PCT/EP2004/053030

Europäisches Patentamt	European Patent Office	Office européen des brevets

19. 11. 2004

REC'D 14 JAN 2005
WIPO PCT

Bescheinigung Certificate Attestation

Die angehefteten Unterlagen stimmen mit der ursprünglich eingereichten Fassung der auf dem nächsten Blatt bezeichneten europäischen Patentanmeldung überein.

The attached documents are exact copies of the European patent application described on the following page, as originally filed.

Les documents fixés à cette attestation sont conformes à la version initialement déposée de la demande de brevet européen spécifiée à la page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

03104280.7

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Offic

Le Président de l'Office européen des brevets
p.o.

R C van Dijk

BEST AVAILABLE COPY

Anmeldung Nr:
Application no.: 03104280.7
Demande no:

Anmeldetag:
Date of filing: 19.11.03
Date de dépôt:

Anmelder/Applicant(s)/Demandeur(s):

CropDesign N.V.
Technologiepark 3
9052 Zwijnaarde-Gent
BELGIQUE

Bezeichnung der Erfindung/Title of the invention/Titre de l'invention:
(Falls die Bezeichnung der Erfindung nicht angegeben ist, siehe Beschreibung.
If no title is shown please refer to the description.
Si aucun titre n'est indiqué se referer à la description.)

Seedy1 sequence for making plants having changed growth characteristics

In Anspruch genommene Prioriät(en) / Priority(ies) claimed /Priorité(s) revendiquée(s)
Staat/Tag/Aktenzeichen/State/Date/File no./Pays/Date/Numéro de dépôt:

Internationale Patentklassifikation/International Patent Classification/
Classification internationale des brevets:

C07K14/00

Am Anmeldetag benannte Vertragstaaten/Contracting states designated at date of filing/Etats contractants désignées lors du dépôt:

AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL
PT RO SE SI SK TR LI

Seedy1 sequence for making plants having changed growth characteristics

The present invention concerns a method for modifying growth characteristics of a plant. More specifically, the present invention concerns a method for modifying growth characteristics by modified expression of a seedy1 nucleic acid and/or by modified levels and/or activity of a seedy1 protein in a plant. The present invention also concerns plants having modified expression of a seedy1 nucleic acid and/or modified levels and/or activity of a seedy1 protein, which plants have modified growth characteristics relative to corresponding wild type plants.

The ever-increasing world population and the dwindling supply of arable land available for agriculture fuel research towards improving the efficiency of agriculture. Conventional means for crop and horticultural improvements utilise selective breeding techniques to identify plants having desirable characteristics. However, such selective breeding techniques have several drawbacks, namely that these techniques are typically labour intensive and result in plants that often contain heterogeneous genetic components that may not always result in the desirable trait being passed on from parent plants. Advances in molecular biology have allowed mankind to modify the germplasm of animals and plants. Genetic engineering of plants entails the isolation and manipulation of genetic material (typically in the form of DNA or RNA) and the subsequent introduction of that genetic material into a plant. Such technology has the capacity to deliver crops or plants having various improved economic, agronomic or horticultural traits. A trait of particular economic interest is yield. Yield is normally defined as the measurable produce of economic value from a crop. This may be defined in terms of quantity and/or quality. Crop yield may not only be increased by combating one or more of stresses to which a crop or plant is typically subjected, but may also be increased by modifying the inherent growth characteristics of a plant. Yield is directly dependent on several growth characteristics, for example, the growth rate, the biomass production, plant architecture, number and size of the organs, (for example, the number of branches, tillers, shoots, flowers), seed production and more.

The ability to influence one or more of the abovementioned growth characteristics, would have many applications in areas such as crop enhancement, plant breeding, production of ornamental plants, aboriculture, horticulture, forestry, production of algae or plants (for example for use as bioreactors, for the production of substances such as pharmaceuticals, antibodies, or vaccines, or for the bioconversion of organic waste or for use as fuel in the case of high-yielding algae and plants).

It has now been found that modification in a plant of the expression of a seedy1 nucleic acid and/or modification of the level and/or activity levels in a plant of a seedy1 protein, gives rise to plants having modified growth characteristics. It has been shown that introduction of a seedy1 nucleic acid effects an increase in above ground biomass, increased total seed weight, increased number of seeds and increased number of tillers. It has been found that the hitherto unknown protein seedy1, is a protein having a coiled coil domain and three conserved motifs as presented in SEQ ID NO's 15, 16 and 17.

Therefore, the present invention provides a method for modifying growth characteristics of a plant, comprising modifying in a plant expression of a nucleic acid encoding a seedy1 protein and/or modifying in a plant level and/or activity of a seedy1 protein, wherein said seedy1 protein comprises in the following order from N-terminus to C-terminus at least two motifs, which motifs are:

(i) a motif having at least 80% sequence identity to the sequence represented by SEQ ID NO 15; and/or

(ii) a motif having at least 80% sequence identity to the sequence represented by SEQ ID NO 16, and/or

(iii) a motif having at least 80% sequence identity to the sequence represented by SEQ ID NO 17 and which is a coiled coil motif; and/or

(iv) a motif having at least 80% sequence identity to the sequence represented by SEQ ID NO 18.

and wherein said growth characteristics are modified relative to the corresponding wild-type plants.

Modifying expression of a seedy1 nucleic acid and/or modifying of the activity and/or levels of a seedy1 protein encompasses modifying expression of a gene and/or modifying acti vity and/or levels of a gene product, namely a polypeptide, in specific cells or tissues. The term "modifying" as used herein means increasing, decreasing or changing in time or place. The modified expression, activity and/or levels of a seedy1 gene or protein are modified compared to expression, activity and/or levels of a seedy1 gene or protein in corresponding wild-type plants. The modified gene expression may result from modified expression levels of an endogenous seedy1 gene and/or may result from modified expression levels of a seedy1 gene previously introduced into a plant. Similarly, modified levels and/or activity of a seedy1 protein may be due to modified expression of an endogenous seedy1 nucleic acid/gene and/or due to modifyed expression of a seedy1 nucleic acid/gene previously introduced into a plant. Modified expression of a gene/nucleic acid and/or modifying activity and/or levels of a gene product/protein may be effected, for example, by chemical means and/or recombinant means.

Therefore there is provided by the present invention, a method for modifying growth characteristics of a plant, comprising modifying seedy1 gene expression and/or modifying seedy1 protein levels and/or seedy1 protein activity, which modification may be effected by recombinant means and/or by chemical means.

Advantageously, modified expression of a seedy1 nucleic acid and/or modified activity and/or levels of a seedy1 protein may be effected by chemical means, i.e. by exogenous application of one or more compounds or elements capable of modifying expression of a seedy1 nucleic acid and/or capable of modifying activity and/or levels of the seedy1 protein. The term "exogenous application" as defined herein is taken to mean the contacting or administering of a suitable compound or element to a plant (as defined herein below, the term "plant" includes plant cell, tissue, organ or to the whole organism). The compound or element may be exogenously applied to a plant in a form suitable for plant uptake (such as through application to the soil for uptake via the roots, or in the case of some plants by applying directly to the leaves, for example by spraying). The exogenous application may take place on wild-type plants or on transgenic plants that have previously been transformed with a seedy1 nucleic acid/gene or another transgene.

Suitable compounds or elements include seedy1 proteins or seedy1 nucleic acids. Alternatively, exogenous application of compounds or elements capable of modifying levels of factors that directly or indirectly activate or inactivate a seedy1 protein will also be suitable in practising the invention. Also included are antibodies that can recognise or mimic the function of seedy1 proteins. Such antibodies may comprise "plantibodies", single chain antibodies, IgG antibodies and heavy chain camel antibodies, as well as fragments thereof. Additionally or alternatively, the resultant effect may also be achieved by the exogenous application of an interacting protein or activator or an inhibitor of the seedy1 gene/gene product. Additionally or alternatively, the compound or element may be a mutagenic substance, such as a chemical selected from any one or more of: N-nitroso-N-ethylurea, ethylene imine, ethyl methanesulphonate and diethyl sulphate. Mutagenesis may also be achieved by exposure to ionising radiation, such as X-rays or gamma-rays or ultraviolet light. Methods for introducing mutations and for testing the effect of mutations (such as by monitoring gene expression and/or protein activity) are well known in the art.

Therefore, according to one aspect of the present invention, there is provided a method for modifying growth characteristics of a plant, comprising exogenous application of one or more

compounds or elements capable of modifying expression of a nucleic acid encoding a seedy1 protein and/or capable of modifying activity and/or levels of a seedy1 protein.

Additionally or alternatively, and according to a preferred embodiment of the present invention, modification of expression of a seedy1 nucleic acid and/or modification of activity and/or levels of a seedy1 protein may be effected by recombinant means. Such recombinant means may comprise a direct and/or indirect approach for modification of expression of a nucleic acid encoding a seedy 1 and/or for modification of the activity and/or levels of a seedy1 protein.

The nucleic acid encoding a seedy one protein or the seedy1 protein as mentioned above may be wild type, i.e. a native or endogenous nucleic acid or protein. Alternatively, it may be a nucleic acid derived from the same or another species, which nucleic acid is introduced as a transgene, for example by transformation. This transgene may be substantially changed from its native form in composition and/or genomic environment through deliberate human manipulation.

An indirect recombinant approach may comprise for example introducing, into a plant, a nucleic acid capable of modifying expression of the gene in question (a seedy1 gene) and/or capable of modifying activity and/or levels of the protein in question (a seedy1 protein). Examples of such nucleic acids to be introduced into a plant are nucleic acids encoding transcription factors or activators or inhibitors that bind to the promoter of a seedy1 gene or that interact with a seedy1 protein. Methods to test these types of interactions and methods for isolating nucleic acids encoding such interactors include yeast one-hybrid or a yeast two-hybrid screens wherein the seedy1 gene/protein is used as a bait.

Also encompassed by an indirect approach for modifying expression of a seedy1 gene and/or activity and/or levels of a seedy1 protein, is the provision of, or the inhibition or stimulation of regulatory sequences that drive expression of the native seedy1 gene or of the seedy1 transgene. Such regulatory sequences may be introduced into a plant. For example, the nucleic acid introduced into the plant is a promoter, capable of driving the expression of an endogenous seedy1 gene.

A further indirect approach for modifying expression of a seedy1 gene and/or for modifying activity and/or levels of a seedy1 protein in a plant, encompasses modified levels in a plant of a factor able to interact with seedy1. Such factors may include ligands of seedy1. Therefore, the present invention provides a method for modifying growth characteristics of a plant, comprising modifying expression of a gene coding for a protein which is a natural ligand of a

seedy1. Furthermore, the present invention also provides a method for modifying grow characteristics of a plant, comprising modifying expression of a gene coding for a protein whi is a natural target/substrate of a seedy1.

A direct and more preferred approach for modifying growth characteristics of a plar comprises introducing into a plant a seedy1 nucleic acid, or a portion thereof or sequence capable of hybridising therewith, which nucleic acid preferably encodes a seedy1 protein or homologue, derivative or active fragment thereof. The nucleic acid may be introduced into plant by, for example, transformation.

Accordingly, the present invention provides a method for modifying growth characteristics of plant, comprising introducing into a plant a nucleic acid capable of modifying expression of nucleic acid encoding a seedy1 protein and/or capable of modifying activity and/or levels of seedy1 protein. Further preferably such nucleic acid is a seedy1 nucleic acid.

As mentioned above the nucleic acid to be used in the methods of the present invention ca be wild type (native or endogenous). Alternatively, the nucleic acid may be derived fro another species, which gene is introduced into the plant as a transgene, for example b transformation. The nucleic acid may thus be derived (either directly or indirectly subsequently modified)) from any source provided that the nucleic acid, when expressed in plant, leads to modified expression of a seedy1 nucleic acid/gene or modified activity and/ levels of a seedy1 protein. The nucleic acid may be isolated from a microbial source, such a bacteria, yeast or fungi, or from a plant, algae, insect, or animal (including human) sourc Preferably, the seedy1 nucleic acid is isolated from a plant. This nucleic acid may b substantially changed from its native form in composition and/or genomic environment throug deliberate human manipulation. The nucleic acid may be isolated from a dicotyledonou species, preferably from the family *Solanaceae*, further preferably from *Nicotiana*. Mor preferably, the nucleic acid is as represented by SEQ ID NO: 1 or a portion thereof or a nuclei acid capable of hybridising therewith or is a nucleic acid encoding an amino acid represente by SEQ ID NO: 2 or a homologue derivative or active fragment thereof, such as a homologu having at least 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80% 85%, 90%, 95% 98%, 99% sequence identity with SEQ ID NO 2.

Advantageously, the methods according to the invention may also be practised using varia nucleic acids and variant amino acids of SEQ ID NO 1 or 2 respectively, such as the variant further defined hereinafter.

The term "seedy1 gene" or "seedy1 nucleic acid" or "nucleic acid encoding a seedy1 protein" are used herein interchangeably.

Taken in a broad context, the term "seedy" protein/nucleic acid also encompasses variant nucleic acids and variant amino acids suitable for practicing the methods according to the invention. Preferably, variant nucleic acids and variant amino acids suitable for practicing the methods according to the invention include those falling within the definition of a "seedy1" meaning a encoding a protein ore being a protein comprising in the following order from N-terminus to C-terminus at least two motifs, which motifs are:

(i) a motif having at least 80% sequence identity to the sequence represented by SEQ ID NO 15; and/or

(ii) a motif having at least 80% sequence identity to the sequence represented by SEQ ID NO 16, and/or

(iii) a motif having at least 80% sequence identity to the sequence represented by SEQ ID NO 17 and which is a coiled coil motif; and/or

(iv) a motif having at least 80% sequence identity to the sequence represented by SEQ ID NO 18.

For determining the presence of these conserved motifs sequence can be aligned with the software such as for example Align X or clustal X, for indication of the conserved residues (see for example Figure 3). Software packages like MEME version 3.0 can also be used to determine motifs in sequences. This software is available from UCSD, SDSC and NBCR at http://meme.sdsc.edu/meme/. For the identification of a coiled coil domain, the software Coils 2.0 can be used. This software is available at http://www.ch.embnet.org/software/COILS_form.html. For the presentation of the motifs in SEQ Id NO 15, 16, 17 and 18, X represents any amino acid.

According to a particular embodiment a seedy 1 protein has at least two, preferably three, most preferably four of the above defined domains.

According to a particular embodiment of the invention, a motif having at least 80% sequence identity to the sequence represented by SEQ ID NO 15 is the core sequence WXNAXXD as represented in SEQ ID NO 15, or at least part of the sequence (P/X)X((V/L/H)(Q/H)(V/I)W(N/X)NA(A/P)(F/C)D wherein

(P/X) preferably is P or is A or T or Q or another amino acid

(V/L/H) preferable is V or L or H

(Q/H) is either Q or H

(V/I) is either V or is T or S or another amino acid

(A/P) is preferable A or is P
(F/C) is preferably F or is C

Alternatively or additionally, according to one embodiment a motif having at least 80% sequence identity to the sequence represented by SEQ ID NO 16, is the core sequence KENXXP as represented in SEQ ID NO 16.

Alternatively or additionally, according to one embodiment a motif having at least 80% sequence identity to the sequence represented by SEQ ID NO 17 and which is a coiled coil motif is the core sequence EX_{1-6}EXXRLXXXLXXLR as represented in SEQ ID NO 17, or is at least part of the sequence (I/V/A)(D/E)XE(I/M)XX(I/V)(E/Q)XE(I/X)XRL(S/X)(S/X)(R/K)LXXLR(L/V/T/I)X(K/Q), wherein

(I/V/A) preferable is I or V or is A
(D/E) is either D or E
(I/M) preferably is I or is M
(I/V) preferably is I or is V
(E/Q) preferably is E or is Q
(I/X) preferably is I or is M or is V or any other amino acid
(S/X) preferably S or is T or any other amino acid
(S/X) preferably is S or isT or L or I or A
(R/K) preferably is R or is K
(L/V/T/I) preferable is L or T or V or I
(K/Q) preferably is K or Q

Alternatively or additionally, according to one embodiment a motif having at least 80% sequence identity to the sequence represented by SEQ ID NO 18, is the core sequence LPXIX_{1-10}RDSGXXKRX_{1-6}K as represented in SEQ ID NO 18, or is at least part of the sequence LP(R/K)I(R/X)(T/I)(M/X)(P/R)XX(D/X)(E/G)(S/T)(P/L)RDSG(C/X)(A/X)KR(V/X)(A/I)(D/E)(L/R)(V/X)(G/A)K, wherein

(R/K) is either R or K
(R/X) is preferably R or is S or K
(T/I) is preferably T or I
(M/X) is preferably M or L or A or V
(P/R) is either P or R
(D/X) is preferably D or is G or T or N
(E/G) is preferably E or is G

(S/T) is preferably S or is T

(P/L) is preferably P or is L

(C/X) is preferably C or is P or A

(A/X) is preferably A or is V or I

(A/I) is preferably A or is I

(D/E) is either D or E

(L/R) is preferably L or is R

(V/X) is preferably V or is Q or N or I

(G/A) is preferably G or is A

According to a further embodiment the nucleic acid has a motif which has at least 80% sequence identity to the sequence RDSGXXKRX$_{1\text{-}6}$K.

Examples of such seedy1 proteins are monocots seedy1 proteins such as represented by SEQ ID NO 4 (rice), SEQ ID NO 8 (sugar cane) and SEQ ID NO 10 (maize) or from dicots SEQ ID NO 2 (tobacco), SEQ ID NO 6 (medicago) or SEQ ID NO 12 (Arabidopsis). The proteins as presented in SEQ ID NO 8 (sugar cane) and SEQ ID NO 10 (Maize) are only partial, but the corresponding full length sequences of the proteins and encoding cDNA, are well within the reach of a person skilled in the art who is now able to perform colony hybridization of a cDNA library, or PCR based on the use of specific primers combined with degenerated primers.

Suitable variant nucleic acid and amino acid sequences useful in practising the method according to the invention, include:

(i) Functional portions of a seedy1 nucleic acid/gene;

(ii) Sequences capable of hybridising with a seedy1 nucleic acid/gene;

(iii) Alternative splice variants of a seedy1 nucleic acid/gene;

(iv) Allelic variants of a seedy1 nucleic acid/gene;

(v) Homologues, derivatives and active fragments of a seedy1 protein;

The term seedy1 nucleic acid/gene, as defined herein, also encompasses a complement of SEQ ID NO 1 and also to corresponding RNA, DNA, cDNA or genomic DNA. The seedy1 may be synthesized in whole or in part, it may be double-strand nucleic acid or single-stranded nucleic acid. Also this term encompasses a variant of the gene due to the degeneracy of the genetic code and variants that are interrupted by one or more intervening sequences.

An example of a variant seedy1 nucleic acid is a functional portion of a seedy1 nucleic acid. The methods according to the invention may advantageously be practised using functional portions of a seedy1. A functional portion refers to a piece of DNA derived or prepared from

an original (larger) DNA molecule, which DNA portion, when introduced and expressed in plant, gives plants having changed development. The portion may comprise many genes, wi or without additional control elements or may contain spacer sequences. The portion may b made by making one or more deletions and/or truncations to the nucleic acid. Techniques fc introducing truncations and deletions into a nucleic acid are well known in the art. Portion suitable for use in the methods according to the invention may readily be determined b following the methods described in the Examples section by simply substituting the sequenc used in the actual Example with the portion to be tested for functionality.

An example of a further variant seedy1 nucleic acid is a sequence that is capable o hybridising to a seedy1 nucleic acid, for example to SEQ ID NO 1, 3, 5, 7, 9 or 11 Advantageously, the methods according to the present invention may also be practised usin sequences capable of hybridising to a coiled coil protein , particularly a seedy1 protein a represented by any one of SEQ ID NO: 2, 4, 6, 8, 10 or 12, which hybridising sequences are preferably those falling within the definition of a "seedy1" as set out herein before. Hybridisin sequences suitable for use in the methods according to the invention may readily be determined for example by following the methods described in the Examples section by simply substituting the sequence used in the actual Example with the hybridising sequence.

The term "hybridisation" as defined herein is a process wherein substantially homologous complementary nucleotide sequences anneal to each other. The hybridisation process can occur entirely in solution, i.e. both complementary nucleic acids are in solution. Tools in molecular biology relying on such a process include the polymerase chain reaction (PCR; and all methods based thereon), subtractive hybridisation, random primer extension, nuclease S1 mapping, primer extension, reverse transcription, cDNA synthesis, differential display of RNAs, and DNA sequence determination. The hybridisation process can also occur with one of the complementary nucleic acids immobilised to a matrix such as magnetic beads, Sepharose beads or any other resin. Tools in molecular biology relying on such a process include the isolation of poly (A+) mRNA. The hybridisation process can furthermore occur with one of the complementary nucleic acids immobilised to a solid support such as a nitro-cellulose or nylon membrane or immobilised by e.g. photolithography to e.g. a siliceous glass support (the latter known as nucleic acid arrays or microarrays or as nucleic acid chips). Tools in molecular biology relying on such a process include RNA and DNA gel blot analysis, colony hybridisation, plaque hybridisation, *in situ* hybridisation and microarray hybridisation. In order to allow hybridisation to occur, the nucleic acid molecules are generally thermally or chemically denatured to melt a double strand into two single strands and/or to remove hairpins or other secondary structures from single stranded nucleic acids. The stringency of hybridisation is

influenced by conditions such as temperature, salt concentration and hybridisation buffer composition. High stringency conditions for hybridisation include high temperature and/or low salt concentration (salts include NaCl and Na_3-citrate) and/or the inclusion of formamide in the hybridisation buffer and/or lowering the concentration of compounds such as SDS (detergent) in the hybridisation buffer and/or exclusion of compounds such as dextran sulphate or polyethylene glycol (promoting molecular crowding) from the hybridisation buffer. Conventional hybridisation conditions are described in, for example, Sambrook (2001) Molecular Cloning: a laboratory manual, 3rd Edition Cold Spring Harbor Laboratory Press, CSH, New York, but the skilled craftsman will appreciate that numerous different hybridisation conditions can be designed in function of the known or the expected homology and/or length of the nucleic acid. Sufficiently low stringency hybridisation conditions are particularly preferred (at least in the first instance) to isolate nucleic acids heterologous to the DNA sequences of the invention defined supra. An example of low stringency conditions is 4-6x SSC / 0.1-0.5% w/v SDS at 37-45°C for 2-3 hours. Depending on the source and concentration of the nucleic acid involved in the hybridisation, alternative conditions of stringency may be employed, such as medium stringency conditions. Examples of medium stringency conditions include 1-4x SSC / 0.25% w/v SDS at ≥ 45°C for 2-3 hours. An example of high stringency conditions includes 0.1-1x SSC / 0.1% w/v SDS at 60°C for 1-3 hours. The skilled man will be aware of various parameters which may be altered during hybridisation and washing and which will either maintain or change the stringency conditions. The stringency conditions may start low and be progressively increased until there is provided a hybridising seedy1 nucleic acid, as defined hereinabove. Elements contributing to heterology include allelism, degeneration of the genetic code and differences in preferred codon usage.

Another example of a variant seedy1 is an alternative splice variant of a seedy1. The methods according to the present invention may also be practised using an alternative splice variant of a seedy1 nucleic acid. The term "alternative splice variant" as used herein encompasses variants of a nucleic acid in which selected introns and/or exons have been excised, replaced or added. Such splice variants may be found in nature or can be manmade using techniques well known in the art. Preferably, the splice variant is a splice variant of the sequence represented by any of SEQ ID NO 1, 3, 5, 7, 9 or 11. Splice variants suitable for use in the methods according to the invention may readily be determined for example by following the methods described in the Examples section by simply substituting the sequence used in the actual Example with the splice variant.

Another example of a variant seedy1 is an allelic variant. Advantageously, the methods according to the present invention may also be practised using allelic variants of a seedy1

nucleic acid, preferably an allelic variant of a sequence represented by any of SEQ ID NO 1, 3, 5, 7, 9 or 11. Allelic variants exist in nature and encompassed within the methods of the present invention is the use of these isolated natural alleles in the methods according to the invention. Allelic variants encompass Single Nucleotide Polymorphisms (SNPs), as well as Small Insertion/Deletion Polymorphisms (INDELs). The size of INDELs is usually less than 100 bp). SNPs and INDELs form the largest set of sequence variants in naturally occurring polymorphic strains of most organisms. Allelic variants suitable for use in the methods according to the invention may readily be determined for example by following the methods described in the Examples section by simply substituting the sequence used in the actual Example with the allelic variant.

Accordingly, the present invention provides a method for modifying growth characteristics of a plant, wherein the seedy1 nucleic acid is a splice variant of a seedy1 nucleic acid or wherein said seedy1 protein is encoded by a splice variant or wherein the seedy1 nucleic acid is an allelic variant of a seedy1 nucleic acid or wherein said seedy1 protein is encoded by an allelic variant.

Examples of variant seedy1 amino acids also in clude homologues, derivatives and active fragments of a seedy1 protein. Advantageously, the methods according to the present invention may also be practised using homologues, derivatives or active fragments of a coiled coil, preferably using homologues, derivatives or active fragments of a seedy1 protein, preferably a seedy1 protein as represented by any one of SEQ ID NO 2, 4, 6, 8, 10 or 12.

"Homologues" of a seedy1 protein encompass peptides, oligopeptides, polypeptides, proteins and enzymes having amino acid substitutions, deletions and/or insertions relative to the unchanged protein in question and having similar biological and functional activity as the unchanged protein from which they are derived. To produce such homologues, amino acids of the protein may be replaced by other amino acids having similar properties (such as similar hydrophobicity, hydrophilicity, antigenicity, propensity to form or break α-helical structures or β-sheet structures). Conservative substitution tables are well known in the art (see for example Creighton (1984) Proteins. W.H. Freeman and Company).

The homologues useful in the methods according to the invention have a percentage identity to any one of SEQ ID NO 2, 4, 6, 8, 10 or 12 equal to value lying between 20% and 99.99%. The homologues useful in the method according to the invention have at least 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, or 50% sequence identity or

similarity (functional identity) to the unchanged protein, alternatively at least 60% sequence identity or similarity to an unchanged protein, alternatively at least 70% sequence identity or similarity to an unchanged protein. Typically, the homologues have at least 75% or 80% sequence identity or similarity to an unchanged protein, preferably at least 85%, 86%, 87%, 88%, 98% sequence identity or similarity, further preferably at least 90%, 91%, 92%, 93%, 94% sequence identity or similarity to an unchanged protein, most preferably at least 95%, 96%, 97%, 98% or 99% sequence identity or similarity to an unchanged protein.
The percentage of identity can be calculated by using an alignment program well known in the art. For example, the percentage of identity can be calculated using the program GAP, or needle (EMBOSS package) or stretcher (EMBOSS package) or the program align X, as a module of the vector NTI suite 5.5 software package, using the standard parameters (for example GAP penalty 5, GAP opening penalty 15, GAP extension penalty 6.6).

The homologues useful in the methods according to the invention are preferably coiled coil proteins, further preferably seedy1 proteins as defined herein above. Homologues suitable for use in the methods according to the invention may readily be determined for example by following the methods described in the Examples section by simply substituting the sequence used in the actual Example with the homologous sequence.

Methods for the search and identification of seedy1 homologues or DNA sequences encoding a seedy1 homologue, would be well within the realm of persons skilled in the art. Such methods, involve screening sequence databases with the sequences as provided by the present invention in SEQ ID NO 1 and 2 or 3 to 10, preferably a computer readable format of the nucleic acids of the present invention. This sequence information is available for example in public databases, that include but are not limited to Genbank (http://www.ncbi.nlm.nih.gov/web/Genbank), the European Molecular Biology Laboratory Nucleic acid Database (EMBL) (http:/w.ebi.ac.uk/ebi-docs/embl-db.html) or versions thereof or the MIPS database (http://mips.gsf.de/). Different search algorithms and software for the alignment and comparison of sequences are well known in the art. Such software includes software include GAP, BESTFIT, BLAST, FASTA and TFASTA. GAP uses the algorithm of Needleman and Wunsch (J. Mol. Biol. 48: 443-453, 1970) to find the alignment of two complete sequences that maximises the number of matches and minimises the number of gaps. The BLAST algorithm calculates percentage sequence identity and performs a statistical analysis of the similarity between the two sequences. The suite of programs referred to as BLAST programs has 5 different implementations: three designed for nucleotide sequence queries (BLASTN, BLASTX, and TBLASTX) and two designed for protein sequence queries (BLASTP and TBLASTN) (Coulson, Trends in Biotechnology: 76-80, 1994; Birren et al.,

GenomeAnalysis, 1: 543, 1997). The software for performing BLAST analysis is publicl available through the National Centre for Biotechnology Information.

Homologues of SEQ ID NO 2 can be found in many different organisms. The closes homologues are found in the plant kingdom. For example, seedy1 nucleic acids were isolat from tobacco (SEQ ID NO 1 and 2), from rice (SEQ ID NO 3) encoding a rice seedy homologue (SEQ ID NO 4), from medicago (SEQ ID NO 5 and 6), from sugar cane (SEQ IL NO 7 and 8), from maize (SEQ ID NO 9 and 10) and from Arabidopsis (SEQ ID NO 11 anc 12). Also EST's from other organisms were found and deposited in Genbank, for example EST from Vitis vinifera (accession number CA816066), from Pinus taeda (accession numbe BM903108), from Saccharus sp. (So variant) (accession numbers CA228193 and CA256020), from Citrus sinsensis (accession number CF833583), Plumbago zeylanica (accession numbe CB817788), from Zea mays (accession number CF637447, AW282224, CD058812, AY108162, CD059048, CF041861, AW067243), from Triticum aestivum (CA727065, BJ264506, BJ259034), from Hordeum vulgare (accession number BU997034, CA727065, CA031127, BQ762011), from Bn (CD817460) from Ga (BG446106, BM360339), from Eschscholzia califronica (CD478368), from Pt (BU821376), and from beta vulgaris (BQ594009).

As more genomes are being sequenced, it is expected that many more seedy1 homologues shall be identifiable.

The above-mentioned analyses for comparing sequences, for calculation of sequence identity and for the search for homologues, is preferentially done with the full-length sequence or with a conserved region of the sequence. Therefore, these analyses can be based on a comparison of certain regions such as conserved domains, motifs or boxes.

The identification of such domains or motifs, would also be well within the realm of a person skilled in the art and involves for example, a computer readable format of the nucleic acids of the present invention, the use of alignment software programs and the use of publicly available information on protein domains, conserved motifs and boxes. This protein domain information is available in the PRODOM (http://www.biochem.ucl.ac.uk/bsm/dbbrowser/jj/prodomsrchjj.html), PIR (http://pir.georgetown.edu/) or pFAM (http://pfam.wustl.edu/) database. Sequence analysis programs designed for motif searching can be used for identification of fragments, regions and conserved domains as mentioned above. Preferred computer programs would include but are not limited to MEME, SIGNALSCAN, and GENESCAN. A MEME algorithm (Version 3.0) can be found in the GCG package; or on the Internet site http://www.sdsc.edu/MEME/meme. SIGNALSCAN version 4.0 information is available on the Internet site

http://biosci.cbs.umn.edu/software/sigscan.html. GENESCAN can be found on the Internet site http://gnomic.stanford.edu/GENESCANW.html.

More particularly preferred seedy1 homologues have the conserved domains as depicted in SEQ ID NO 15, 16 and 17, or motifs that are 80% identical to these motifs. Also preferred seedy1 homologues have a coiled coil domain, preferably located in the N-terminal half of the protein, more preferably between amino acid position 25 to 250, more preferably between position 50 and 150. A coiled coil domain is determining the folding of the protein.

Two special forms of homology, orthologous and paralogous, are evolutionary concepts used to describe ancestral relationships of genes. The term "paralogous" relates to gene-duplications within the genome of a species. A "paralogue" is therefore duplication. The term "orthologous" relates to homologous genes in different organisms due to ancestral relationship and the formation of different species. An orthologue is thus a "speciation". The term "homologues" as used herein also encompasses paralogues and orthologues and are useful proteins in the methods according to the invention.

Another variant of seedy1 useful in the methods of the present invention is a derivative of seedy1. The term "derivatives" refers to peptides, oligopeptides, polypeptides, proteins and enzymes which may comprise substitutions, deletions or additions of naturally and non-naturally occurring amino acid residues compared to the amino acid sequence of a naturally-occurring form of the protein, for example, as presented in SEQ ID NO: 2. "Derivatives" of a seedy1 protein encompass peptides, oligopeptides, polypeptides, proteins and enzymes which may comprise naturally occurring changed, glycosylated, acylated or non-naturally occurring amino acid residues compared to the amino acid sequence of a naturally-occurring form of the polypeptide. A derivative may also comprise one or more non-amino acid substituents compared to the amino acid sequence from which it is derived, for example a reporter molecule or other ligand, covalently or non-covalently bound to the amino acid sequence such as, for example, a reporter molecule which is bound to facilitate its detection, and non-naturally occurring amino acid residues relative to the amino acid sequence of a naturally-occurring protein.

"Substitutional variants" of a protein are those in which at least one residue in an amino acid sequence has been removed and a different residue inserted in its place. Amino acid substitutions are typically of single residues, but may be clustered depending upon functional constraints placed upon the polypeptide; insertions will usually be of the order of about 1 to 10

amino acid residues, and deletions will range from about 1 to 20 residues. Preferably, amino acid substitutions comprise conservative amino acid substitutions.

"Insertional variants" of a protein are those in which one or more amino acid residues are introduced into a predetermined site in a protein. Insertions can comprise amino-terminal and/or carboxy-terminal fusions as well as intra-sequence insertions of single or multiple amino acids. Generally, insertions within the amino acid sequence will be smaller than amino- or carboxy-terminal fusions, of the order of about 1 to 10 residues. Examples of amino- or carboxy-terminal fusion proteins or peptides include the binding domain or activation domain of a transcriptional activator as used in the yeast two-hybrid system, phage coat proteins, (histidine)6-tag, glutathione S-transferase-tag, protein A, maltose-binding protein, dihydrofolate reductase, Tag·100 epitope, c-myc epitope, FLAG®-epitope, lacZ, CMP (calmodulin-binding peptide), HA epitope, protein C epitope and VSV epitope.

"Deletion variants" of a protein are characterised by the removal of one or more amino acids from the protein. Amino acid variants of a protein may readily be made using peptide synthetic techniques well known in the art, such as solid phase peptide synthesis and the like, or by recombinant DNA manipulations. Methods for the manipulation of DNA sequences to produce substitution, insertion or deletion variants of a protein are well known in the art. For example, techniques for making substitution mutations at predetermined sites in DNA are well known to those skilled in the art and include M13 mutagenesis, T7-Gen in vitro mutagenesis (USB, Cleveland, OH), QuickChange Site Directed mutagenesis (Stratagene, San Diego, CA), PCR-mediated site-directed mutagenesis or other site-directed mutagenesis protocols.

Another variant of seedy1 useful in the methods of the present invention is an active fragment of seedy1. "Active fragments" of a seedy1 protein encompasses contiguous amino acid residues of a seedy1 protein, which residues retain similar biological and/or functional activity to the naturally occurring protein. For example, useful fragments comprise at least 10 contiguous amino acid residues of a seedy1 protein. Other preferred fragments are fragments of the seedy1 protein starting at the second or third or further internal methionin residues. These fragments originate from protein translation, starting at internal ATG codons.

In the present invention a new category of proteins and encoding genes is disclosed. Surprisingly, altering the availability of such sequences is a plant changes the growth characteristics of the plant. More particularly the seed quantity and quality is improved in these plants. Therefore, these sequences useful for the methods of the present invention are named "seedy" sequences.

As the type of proteins defined herein as seedy1 proteins was not known, the present invention provides for an isolated nucleic acid encoding at least part of a seedy1 protein, wherein said seedy1 protein comprises in the following order from N-terminus to C-terminus at least two motifs, which motifs are:

(i) a motif having at least 80% sequence identity to the sequence represented by SEQ ID NO 15; and/or

(ii) a motif having at least 80% sequence identity to the sequence represented by SEQ ID NO 16, and/or

(iii) a motif having at least 80% sequence identity to the sequence represented by SEQ ID NO 17 and which is a coiled coil motif; and/or

(iv) a motif having at least 80% sequence identity to the sequence represented by SEQ ID NO 18.

In the present invention, a new type of gene has been isolated from tobacco which gene encodes a new type of protein. Surprisingly it has been found that this gene and/or protein is capable of modifying growth characteristics of a plant. Therefore, the present invention provides an isolated nucleic acid selected from the group comprising,

(i) a nucleic acid represented by any of SEQ ID NO: 1, 5 or 7 the complement strand thereof;

(ii) a nucleic acid encoding an amino acid sequence represented by SEQ ID NO: 2, 4, 6, 8 or 10 or a homologue, derivative or active fragment of the above mentioned sequences;

(iii) a nucleic acid capable of hybridising with a nucleic acid of (i) or (ii) above, which hybridising sequence preferably encodes a protein having seedy1 protein activity;

(iv) a nucleic acid which is degenerate from any one of the nucleic acids of (i) to (iii) above as a results of the genetic code;

(v) a nucleic acid which is an allelic variant of any one of the nucleic acids of (i) to (iv);

(vi) a nucleic acid which is an alternative splice variant of any one of the nucleic acids of (i) to (v);

(vii) a nucleic acid encoding a protein which has at least 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% sequence identity to any one or more from the sequences defined in (i) to (vi), which protein preferably encodes a protein having seedy1 activity;

(viii) a portion of a nucleic acid according to any of (i) to (vii) above, which portio preferably encodes a protein having seedy1 activity.

According to another embodiment of the present invention, there is provided an isolate nucleic acid as mentioned herein above, wherein said isolated nucleic acid is not a nuclei acid selected from the group comprising: the rice cDNA as deposited under Genbanl accession number AK063941 (SEQ ID NO 3), a *Medicago* BAC clone deposited as AC144618 or AC139356 or AC144482 or AC135566, the *Arabidopsis* cDNA deposited under AL61572 (SEQ ID NO 11), or the *Zea mays* EST deposited under AY108162 (SEQ ID NO 9). Furthe the present invention described for the first time a novel type of protein and therefore, within the scope of the present invention is an isolated seedy1 protein comprising in the following order from N-terminus to C-terminus at least two motifs, which motifs are:

(i) a motif having at least 80% sequence identity to the sequence represented by SEQ ID NO 15; and/or

(ii) a motif having at least 80% sequence identity to the sequence represented by SEQ ID NO 16, and/or

(iii) a motif having at least 80% sequence identity to the sequence represented by SEQ ID NO 17 and which is a coiled coil motif; and/or

(iv) a motif having at least 80% sequence identity to the sequence represented by SEQ ID NO 18.

Further, the present invention provides, an isolated seedy1 protein, comprising

a. a polypeptide with an amino acid sequence as presented in any one of SEQ ID NO 2, 4, 6, 8 or 10;

b. a polypeptide with an amino acid sequence which has at least 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% sequence identity to any one or more of the amino acid sequence as described in (a)

c. a polypeptide which is a homologue, a derivative, an immunologically active and/or functional fragment of a protein as defined in any of (a) or (b).

According to another embodiment of the present invention, there is provided an isolated seedy1 protein as mentioned herein above, wherein said protein is not the *Arabisopsis* protein as deposited in Genbank under the accession number AL161572 (SEQ ID NO 12).

According to a preferred aspect of the present invention, enhanced or increased expression of a seedy1 nucleic acid in a plant or plant part is envisaged. Methods for obtaining increased expression of genes or gene products are well documented in the art and include, for example, overexpression driven by a (strong) promoter, the use of transcription enhancers or translation enhancers. The term overexpression as used herein means any form of expression that is additional to the original wild-type expression level. Preferably the nucleic acid to be introduced into the plant and/or the nucleic acid that is to be overexpressed in the plants is in the sense direction with respect to the promoter to which it is operably linked. Alternatively, selection of better performing alleles of the wild-type seedy1 nucleic acid can be achieved via plant breeding techniques.

Accordingly, a preferred embodiment of the present invention provides a method to modify growthe characteristics in a plant, comprising introducing, into a plant, a nucleic acid capable of modifying expression of a seedy1 gene and/or capable of modifying activity and/or level of a seedy1 protein in the sense orientation relative to control element to which it is operably linked.

Alternatively and/or additionally, increased expression of a seedy1 encoding gene or increased activities and/or levels of a seedy1 protein in a plant cell, is achieved by mutagenesis. For example these mutations can be responsible for the changed control of the seedy1 gene, resulting in more expression of the gene, relative to the wild-type gene. Mutations can also cause conformational changes in a protein, resulting in more activity and/or levels of the seedy1 protein.

Additionally, these is envisaged by the present invention a method for modifying growth characteristics comprising downregulation of expression of a seedy1 gene or downregulation of levels and/or activity of a seedy1 protein. These methods are particularly useful to modify the size or quantity or architecture of plants particular plant organs. According to a particular embodiment of the present invention, plants are decreased in size or plant organs are decreased in size or number, or the plant architecture is changed, for example there is less branching or less branching of the inflorescence. Therefore, according to a further aspect of the invention, decreased expression of a seedy1 nucleic acid or decreased activity and/or level of a seedy1 is envisaged.

Examples of decreasing or downregulation of expression are well documented in the art and include, for example, downregulation of expression by anti-sense techniques, RNAi techniques, small interference RNAs (siRNAs), microRNA (miRNA), etc. Therefore according to a particular aspect of the invention, there is provided a method for modifying growth characteristics of plants, including technologies that are based on for example the synthesis of antisense transcripts, complementary to the mRNA of a seedy1 gene.

Another method for downregulation of gene expression or gene silencing comprises use o ribozymes, for example as described in WO9400012 (Atkins et al.), WO9503404 (Lenee et al.) WO0000619 (Nikolau et al.), WO9713865 (Ulvskov et al.) and WO9738116 (Scott et al.).

Gene silencing may also be achieved by insertion mutagenesis (for example, T-DNA insertion or transposon insertion) or by gene silencing strategies as described among others in the documents WO9836083 (Baulcombe and Angell), WO9853083 (Grierson et al.), WO9915682 (Baulcombe et al.) or WO9953050 (Waterhouse et al.).

Expression of an endogenous gene may also be reduced if the endogenous gene contains a mutation. Such a mutant gene may be isolated and introduced into the same or different plant species in order to obtain plants having modified growth characteristics. Also dominant negative mutants of a seedy1 nucleic acid can be introduced in the cell to decrease the level/and or activity of the endogenous Seedy1 protein.

Other methods to decrease the expression of a seedy1 nucleic acid and/or activity and/or level of seedy1 proteins in a cell encompass for example the mechanisms of transcriptional gene silencing, such as the methylation of the seedy1 promoter.

Another mechanism to downregulate levels and/or activity of a seedy1 protein in a plant encompasses the mechanism of co-suppression. Modifying gene expression (whether by a direct or indirect approach) encompasses changed transcript levels of that gene. Changed transcript levels can be sufficient to induce certain phenotypic effects, for example via the mechanism of cosuppression. Here the overall effect of expression of a transgene is that there is less activity in the cell of the protein encoded by a native gene having homology to the introduced transgene. Cosuppression is accomplished by the addition of coding sequences or parts thereof in a sense orientation into the cell. Therefore, according to one aspect of the present invention, the development of a plant may be changed by introducing into a plant an additional copy (in full or in part) of a seedy1 gene already present in a host plant. The additional gene may silence the endogenous gene, giving rise to a phenomenon known as co-suppression.

Genetic constructs aimed at silencing gene expression may comprise the seedy1 nucleotide sequence or one at least a portion thereof in a sense and/or antisense orientation relative to the promoter sequence. Preferably the portions comprises at least 21 contiguous nucleic acid of a sequence to be downregulated. Also, sense or antisense copies of at least part of the

endogenous gene in the form of direct or inverted repeats may be utilised in the methods according to the invention. The development of plants may also be changed by introducing into a plant at least part of an antisense version of the nucleotide sequence represented, for example, by SEQ ID NO: 1. It should be clear that part of the nucleic acid (a portion) could also achieve the desired result. Homologous anti-sense genes are preferred, homologous genes being plant genes, preferably plant genes from the same plant species in which the silencing construct is introduced.

The expression of a seedy1 gene can be investigated by Northern or Southern blot analysis of cell extracts. The levels of seedy1 protein in the cell can be investigated via Western blot analysis of cell extracts.

According to a further embodiment of the present invention, genetic constructs and vectors to facilitate introduction and/or expression of the nucleotide sequences useful in the methods according to the invention are provided. Therefore, according to the further embodiment, the present invention provides a genetic construct comprising:

(i) a at least part of a nucleic acid of the present invention as described above or at least a part of a nucleic acid encoding a seedy1 protein as defined above;

(ii) one or more control sequences capable of regulating expression of the nucleic acid of (i); and optionally

(iii) a transcription termination sequence.

According to methods of the present invention, such a genetic construct is introduced into a plant or plant part.

Constructs useful in the methods according to the present invention may be constructed using recombinant DNA technology well known to persons skilled in the art. The gene constructs may be inserted into vectors, which may be commercially available, suitable for transforming into plants and suitable for expression of the gene of interest in the transformed cells.
The genetic construct can be an expression vector wherein said nucleic acid is operably linked to one or more control sequences allowing expression in prokaryotic and/or eukaryotic host cells.

The nucleic acid according to (i) is advantageously any of the aforementioned nucleic acids, preferably any of the above mentioned variants of a seedy1 nucleic acid, most preferably a seedy1 nucleic acid, such as for example one of the nucleic acids of SEQ ID NO 1, 3, 5, 7, 9 or

11. The construct sequence of (ii) is preferably a seed-preferred promoter, for example a prolamin promoter

According to a preferred embodiment of the invention, the genetic construct is an expression vector designed to overexpress the nucleic acid. A preferred nucleic acid (i) in this overexpression vector is the sequence represented by SEQ ID NO 1, 3, 5, 7, 9 or 11 or a portion thereof or sequences capable of hybridising therewith or a nucleic acid encoding a sequence represented by SEQ ID NO 2, 4, 6, 8, 10 or 12 or a homologue, derivative or active fragment thereof. Preferably, this nucleic acid is cloned in the sense orientation relative to the control sequence.

The methods according to the present invention may also be practised by introducing into a plant at least a part of a (natural or artificial) chromosome (such as a Bacterial Artificial Chromosome (BAC)), which chromosome contains at least a seedy1 nucleic acid, optionally together with one or more related gene family members. Therefore, according to a further aspect of the present invention, there is provided a method for modifying growth characteristics of a plant by introducing into a plant at least a part of a chromosome comprising at least a seedy1 gene/nucleic, which seedy1 nucleic is preferably one represented by any one of SEQ ID NO 1, 3, 5, 7, 9 or 11.

Plants are transformed with a vector comprising the sequence of interest (i.e., the nucleic acid capable of modifying expression of seedy1 nucleic acid), which sequence is operably linked to one or more control sequences (at least a promoter). The terms "regulatory element", "control sequence" are all used herein interchangeably and are to be taken in a broad context to refer to regulatory nucleic acids capable of effecting expression of the sequences to which they are ligated (i .e. operably linked). Encompassed by the aforementioned terms are promoters. A "Promoter" encompasses transcriptional regulatory sequences derived from a classical eukaryotic genomic gene (including the TATA box which is required for accurate transcription initiation, with or without a CCAAT box sequence) and additional regulatory elements (i.e. upstream activating sequences, enhancers and silencers) which alter gene expression in response to developmental and/or external stimuli, or in a tissue-specific manner. Also included within the term is a transcriptional regulatory sequence of a classical prokaryotic gene, in which case it may include a –35 box sequence and/or –10 box transcriptional regulatory sequences. The term "regulatory element" also encompasses a synthetic fusion molecule or derivative which confers, activates or enhances expression of a nucleic acid molecule in a cell, tissue or organ. The term "operably linked" as used herein refers to a functional linkage between the promoter sequence and the gene of interest, such that the promoter sequence is able to initiate transcription of the gene of interest.

Advantageously, any type of promoter may be used to drive expression of the nucleic acid depending on the desired outcome. For example, a meristem-specific promoter, such as the rnr (ribonucleotide reductase), cdc2a promoter and the cyc07 promoter. Also seed -specific promoter, such as p2S2, pPROLAMIN, pOLEOSIN could be selected. An aleurone-specific promoter may be selected. An inflorescence-specific promoter, such as pLEAFY, may also be utilised. To produce male-sterile plants one would need an anther specific promoter. One could also choose a petal-specific promoter. If the desired outcome would be to modify plant growth characteristics in particular organs, then the choice of the promoter would depend on the organ to be changed. For example, use of a root-specific promoter would lead to phenotypic alteration of the root. This would be particularly important where it is the root itself that is the desired end product; such crops include sugar beet, turnip, carrot, and potato. A fruit-specific promoter may be used to modify, for example, the strength of the outer skin of the fruit or to increase the size of the fruit. A green tissue-specific promoter may be used to influence the phenotype pf the leaf. A cell wall-specific promoter may be used to increase the rigidity of the cell wall, thereby increasing resistance of the cell. An anther-specific promoter may be used to produce male-sterile plants. A vascular-specific promoter may be used to increase transport from leaves to seeds. A nodule-specific promoter may be used to increase the nitrogen fixing capabilities of a plant, thereby increasing the nutrient levels in a plant. A stress-inducible promoter may also be used to drive expression of a nucleic acid during conditions of stress. A stress inducible promoter such as the water stress induced promoter WSI18, the drought stress induced Trg-31 promoter, the ABA related promoter rab21 or any other promoter which is induced under a particular stress condition such as temperature stress (cold, freezing, heat) or osmotic stress, or drought stress or oxidative stress or biotic stress can be used to drive expression of a seedy1 gene.

Preferably, the nucleic acid capable of modifying expression of a seedy1 gene is operably linked to a plant promoter, preferably a tissue preferred promoter. The term "tissue-preferred" as defined herein refers to a promoter that is expressed predominantly in at least one tissue or organ. Preferably, the tissue-preferred promoter is a seed-preferred promoter or a seed-specific promoter, further preferably an endosperm-specific promoter, more preferably a promoter isolated from a gene encoding a seed-storage protein most preferably a promoter isolated from a prolamin gene, such as for example a rice prolamin promoter as presented in SEQ ID NO 14 or a promoter of similar strength and/or a promoter with a similar expression pattern as the rice prolamin promoter. Similar strength and/or similar expression pattern can be analysed for example by coupling the promoters to a reporter gene and check the function of the reporter gene in tissues of the plant. One suitable reporter gene is beta-glucuronidase and

the colorimetric GUS staining to visualize the beta-glucuronidase activity in a plant tissue i well known to a person skilled in the art.

Examples of preferred seed-specific promoters and other tissue-specific promoters are presented in Table I, which promoters or derivatives thereof are useful in performing the methods of the present invention. Accordingly, genetic constructs comprising at least part of a seedy1 nucleic acid and at least part of a promoter from table I, preferably, wherein said parts are operably linked, is also provided by the present invention.

TABLE I

EXEMPLARY SEED-PREFERRED PROMOTERS FOR USE IN THE PERFORMANCE O THE PRESENT INVENTION		
GENE SOURCE	**EXPRESSION PATTERN**	**REFERENCE**
seed-specific genes	seed	Simon, *et al., Plant Mol. Biol. 5:* 191 1985; Scofield, *et al., J. Biol. Chem 262:* 12202, 1987.; Baszczynski, *et al. Plant Mol. Biol. 14:* 633, 1990.
Brazil Nut albumin	seed	Pearson, et al., Plant Mol. Biol. 18: 235-245, 1992.
legumin	seed	Ellis, et al., Plant Mol. Biol. 10: 203-214, 1988.
glutelin (rice)	seed	Takaiwa, et al., Mol. Gen. Genet. 208: 15-22, 1986; Takaiwa, et al., FEBS Letts. 221: 43-47, 1987.
zein	seed	Matzke et al Plant Mol Biol, 14(3):323-32 1990
napA	seed	Stalberg, *et al, Planta 199:* 515 -519, 1996.
wheat LMW and HMW glutenin-1	endosperm	Mol Gen Genet 216:81-90, 1989; NAR 17:461-2, 1989
wheat SPA	seed	Albani *et al,* Plant Cell, 9: 171-184, 1997
wheat α, β, γ-gliadins	endosperm	EMBO 3:1409-15, 1984
barley *ltr1* promoter	endosperm	
barley B1, C, D, hordein	endosperm	Theor Appl Gen 98:1253-62, 1999; Plant J 4:343-55, 1993; Mol Gen Genet 250:750-60, 1996
barley DOF	endosperm	Mena *et al,* The Plant Journal, 116(1): 53-62, 1998
blz2	endosperm	EP99106056.7

synthetic promoter	endosperm	Vicente-Carbajosa *et al., Plant J. 13:* 629-640, 1998.
rice prolamin NRP33	endosperm	Wu *et al,* Plant Cell Physiology 39(8) 885-889, 1998
rice α-globulin Glb-1	endosperm	Wu *et al,* Plant Cell Physiology 39(8) 885-889, 1998
rice OSH1	embryo	Sato *et al,* Proc. Natl. Acad. Sci. USA, 93: 8117-8122, 1996
rice α-globulin REB/OHP-1	endosperm	Nakase *et al.* Plant Mol. Biol. 33: 513-522, 1997
rice ADP-glucose PP	endosperm	Trans Res 6:157-68, 1997
maize ESR gene family	endosperm	Plant J 12:235-46, 1997
sorgum γ-kafirin	endosperm	PMB 32:1029-35, 1996
KNOX	embryo	Postma-Haarsma *et al,* Plant Mol. Biol. 39:257-71, 1999
rice oleosin	embryo and aleuron	Wu *et at,* J. Biochem., 123:386, 1998
sunflower oleosin	seed (embryo and dry seed)	Cummins, *et al.,* Plant Mol. Biol. 19: 873-876, 1992

Metallothionein Mte, PRO0001	transfer layer of embryo + calli
putative beta-amylase, PRO0005	transfer layer of embryo
putative cellulose synthase, PRO0009	weak in roots
lipase (putative), PRO0012	
transferase (putative), PRO0014	
peptidyl prolyl cis-trans isomerase (putative), PRO0016	
Unknown, PRO0019	
prp protein (putative), PRO0020	
noduline (putative), PRO0029	
proteinase inhibitor Rgpi9, PRO0058	seed
beta expansine EXPB9, PRO0061	weak in young flowers
structural protein, PRO0063	young tissues+calli+embryo
xylosidase (putative), PRO0069	
prolamine 10 Kda, PRO0075	strong in endosperm
allergen RA2, PRO0076	strong in endosperm
prolamine RP7, PRO0077	strong in endosperm
CBP80, PRO0078	
starch branching enzyme I, PRO0079	
Metallothioneine-like ML2, PRO0080	transfer layer of embryo + calli
putative caffeoyl-CoA 3-O-methyltransferase, PRO0081	shoot
prolamine RM9, PRO0087	strong in endosperm

prolamine RP6, PRO0090	strong endosperm
prolamine RP5, PRO0091	strong in endosperm
allergen RA5, PRO0092	
putative methionine aminopeptidase, PRO0095	embryo
ras-related GTP binding protein, PRO0098	
beta expansine EXPB1, PRO0104	
Glycine rich protein, PRO0105	
metallothionein like protein (putative), PRO0108	
metallothioneine (putative), PRO0109	
RCc3, PRO0110	strong root
uclacyanin 3-like protein, PRO0111	weak discrimination center / shoot meristem
26S proteasome regulatory particle non-ATPase subunit 11, PRO0116	very weak meristem specific
putative 40S ribosomal protein, PRO0117	weak in endosperm
chlorophyll a/b-binding protein presursor (Cab27), PRO0122	very weak in shoot
putative protochlorophyllide reductase, PRO0123	strong leaves
metallothionein RiCMT, PRO0126	strong discrimination center / shoot meristem
GOS2, PRO0129	strong constitutive
GOS9, PRO0131	
chitinase Cht-3, PRO0133	very weak meristem specific
alpha-globulin, PRO0135	strong in endosperm
alanine aminotransferase, PRO0136	weak in endosperm
cyclin A2, PRO0138	
Cyclin D2, PRO0139	
Cyclin D3, PRO0140	
cyclophyllin 2, PRO0141	shoot and seed
sucrose synthase SS1 (barley), PRO0146	medium constitutive
trypsin inhibitor ITR1 (barley), PRO0147	weak in endosperm
ubiquitine 2 with intron, PRO0149	strong constitutive
WSI18, PRO0151	embryo + stress
HVA22 homologue (putative), PRO0156	
EL2, PRO0157	
Aquaporine, PRO0169	medium constitutive in young plants
High mobility group protein, PRO0170	strong constitutive
reversibly glycosylated protein RGP1, PRO0171	weak constitutive
cytosolic MDH, PRO0173	shoot
RAB21, PRO0175	embryo + stress

CDPK7, PRO0176	
Cdc2-1, PRO0177	very weak in meristem
sucrose synthase 3, PRO0197	
OsVP1, PRO0198	
OSH1, PRO0200	very weak in young plant meristem
putative chlorophyllase, PRO0208	
OsNRT1, PRO0210	
EXP3, PRO0211	
phosphate transporter OjPT1, PRO0216	
oleosin 18kd, PRO0218	aleurone + embryo
ubiquitine 2 without intron, PRO0219	
RFL, PRO0220	
maize UBI delta intron, PRO0221	
glutelin-1, PRO0223	
fragment of prolamin RP6 promoter, PRO0224	
4xABRE, PRO0225	
glutelin OSGLUA3, PRO0226	
BLZ-2_**short** (barley), PRO0227	
BLZ-2_**long** (barley), PRO0228	

Optionally, one or more terminator sequences may also be used in the construct introduced into a plant. The term "terminator" encompasses a control sequence which is a DNA sequence at the end of a transcriptional unit which signals 3' processing and polyadenylation of a primary transcript and termination of transcription. Additional regulatory elements may include transcriptional as well as translational enhancers. Those skilled in the art will be aware of terminator and enhancer sequences, which may be suitable for use in performing the invention. Such sequences would be known or may readily be obtained by a person skilled in the art.

The genetic constructs of the invention may further include an origin of replication sequence which is required for maintenance and/or replication in a specific cell type. One example is when a genetic construct is required to be maintained in a bacterial cell as an episomal genetic element (e.g. plasmid or cosmid molecule). Preferred origins of replication include, but are not limited to, the f1-ori and colE1.

The genetic construct may optionally comprise a selectable marker gene. As used herein, the term "selectable marker gene" includes any gene, which confers a phenotype on a cell in which it is expressed to facilitate the identification and/or selection of cells which are

transfected or transformed with a nucleic acid construct of the invention. Suitable markers ma be selected from markers that confer antibiotic or herbicide resistance, that introduce a nev metabolic trait or that allow visual selection. Examples of selectable marker genes includ genes conferring resistance to antibiotics (such as nptII encoding neomycir phosphotransferase capable of phosphorylating neomycin and kanamycin, or hpt encodin hygromycin phosphotransferase capable of phosphorylating hygromycin), to herbicides (fo example bar which provides resistance to Basta; aroA or gox providing resistance agains glyphosate), or genes that provide a metabolic trait (such as manA that allows plants to use mannose as sole carbon source). Visual marker genes result in the formation of colour (fo example beta-glucuronidase, GUS), luminescence (such as luciferase) or fluorescence (Green Fluorescent Protein, GFP, and derivatives thereof). Further examples of suitable selectable marker genes include the ampicillin resistance (Ampr), tetracycline resistance gene (Tcr), bacterial kanamycin resistance gene (Kanr), phosphinothricin resistance gene, and the chloramphenicol acetyltransferase (CAT) gene, amongst others

In a preferred embodiment, the genetic construct as mentioned above, comprises at least part of a rice prolamin promoter operably linked to a seedy1 nucleic acid in the sense orientation. An example of such an expression cassette, further comprising a terminator sequence s presented in SEQ ID NO 13.

Therefore the present invention provides an isolated genetic construct, comprising a nucleic acid selected from the group comprising:

a. A nucleic acid having at least a part of a seed-preferred promoter and at least a part of the nucleic acid as defined herein above 2; and
b. A nucleic acid as presented in SEQ ID NO 13, or the complementary strand thereof;
c. a nucleic acid which is degenerated as a result of the genetic code from any of the nucleic acids of (a) or (b);
d. a nucleic acids which is an allelic variant of any of the nucleic acids of (a) or (b);
e. a nucleic acid which hybridizes to any of the nucleic acids of (a) or (b).

According to a further embodiment of the present invention, there is provided a method for the production of a plant having modified growth characteristics, comprising modifying expression and or activity and/or levels in a plant of a seedy1 nucleic acid or seedy1 protein.

According to a particular embodiment, the present invention provides a method for the production of a transgenic plant having modified growth characteristics, which method comprises:

(i) introducing into a plant or plant part a nucleic acid capable of modifying expression of a seedy1 gene and/or capable of modifying the activity and/or levels of a seedy1 protein;

(ii) cultivating the plant cell under conditions promoting regeneration and mature plant growth.

The nucleic acid of (i) may advantageously be any of the aforementioned nucleic acids, preferably a nucleic acid encoding a coiled coil protein, more preferably encoding a seedy1 most preferably a seedy1 nucleic acid according to SEQ ID NO 1, 3, 5, 7, 9, or 11.

Preferably in the above mentioned method, the nucleic acid of (i) is overexpressed in the plant, is in the sense direction and/or is driven by a seed-preferred promoter, such as the prolamin promoter.

The protein itself and/or the nucleic acid itself may be introduced directly into a plant cell or into the plant itself (including introduction into a tissue, organ or any other part of the plant). According to a preferred feature of the present invention, the nucleic acid is preferably introduced into a plant by transformation.

The term "transformation" as referred to herein encompasses the transfer of an exogenous polynucleotide into a host cell, irrespective of the method used for transfer. Plant tissue capable of subsequent clonal propagation, whether by organogenesis or embryogenesis, may be transformed with a genetic construct of the present invention and a whole plant regenerated therefrom. The particular tissue chosen will vary depending on the clonal propagation systems available for, and best suited to, the particular species being transformed. Exemplary tissue targets include leaf disks, pollen, embryos, cotyledons, hypocotyls, megagametophytes, callus tissue, existing meristematic tissue (e.g., apical meristem, axillary buds, and root meristems), and induced meristem tissue (e.g. cotyledon meristem and hypocotyl meristem). The polynucleotide may be transiently or stably introduced into a host cell and may be maintained non-integrated, for example, as a plasmid. Alternatively and preferably, the transgene may be stably integrated into the host genome. The resulting transformed plant cell can then be used to regenerate a transformed plant in a manner known to persons skilled in the art.

Transformation of a plant species is now a fairly routine technique. Advantageously, any of several transformation methods may be used to introduce the gene of interest into a suitable ancestor cell. Transformation methods include the use of liposomes, electroporation, chemicals that increase free DNA uptake, injection of the DNA directly into the plant, particle gun bombardment, transformation using viruses or pollen and microprojection. Methods may be selected from the calcium/polyethylene glycol method for protoplasts (Krens, F.A. et al.,

1882, Nature 296, 72-74; Negrutiu I. et al., June 1987, Plant Mol. Biol. 8, 363-373 electroporation of protoplasts (Shillito R.D. et al., 1985 Bio/Technol 3, 1099-1102 microinjection into plant material (Crossway A. et al., 1986, Mol. Gen Genet 202, 179-185; DNA or RNA-coated particle bombardment (Klein T.M. et al., 1987, Nature 327, 70) infectio with (non-integrative) viruses and the like.

Transgenic rice plants expressing a seedy1 gene are preferably produced via *Agrobacterium* mediated transformation using any of the well known methods for rice transformation, such as described in any of the following: published European patent application EP 1198985 A1 Aldemita and Hodges (Planta, 199, 612-617, 1996); Chan *et al.* (Plant Mol. Biol. 22 (3) 491-506, 1993), Hiei *et al.* (Plant J. 6 (2) 271-282, 1994), which disclosures are incorporated by reference herein as if fully set forth. In the case of corn transformation, the preferred method is as described in either Ishida *et al.* (Nat. Biotechnol. 1996 Jun; 14(6): 745-50) or Frame *et al.* (Plant Physiol. 2002 May; 129(1): 13-22), which disclosures are incorporated by reference herein as if fully set forth.

Generally after transformation, plant cells or cell groupings are selected for the presence of one or more markers which are encoded by plant-expressible genes co-transferred with the gene of interest, following which the transformed material is regenerated into a whole plant.

Following DNA transfer and regeneration, putatively transformed plants may be evaluated, for instance using Southern analysis, for the presence of the gene of interest, copy number and/or genomic organisation. Alternatively or additionally, expression levels of the newly introduced DNA may be monitored using Northern and/or Western analysis, both techniques being well known to persons having ordinary skill in the art.

The generated transformed plants may be propagated by a variety of means, such as by clonal propagation or classical breeding techniques. For example, a first generation (or T1) transformed plant may be selfed to give homozygous second generation (or T2) transformants, and the T2 plants further propagated through classical breeding techniques.

The generated transformed organisms may take a variety of forms. For example, they may be chimeras of transformed cells and non-transformed cells; clonal transformants (e.g., all cells transformed to contain the expression cassette); grafts of transformed and untransformed tissues (e.g., in plants, a transformed rootstock grafted to an untransformed scion).

The present invention also encompasses plants obtainable by the methods according to the present invention. The present invention therefore provides plants obtainable by the method

according to the present invention, which plants have modified growth characteristocs, when compared to the wild-type plants.

The present invention clearly extends to any plant cell or plant produced by any of the methods described herein, and to all plant parts and propagules thereof. The present invention extends further to encompass the progeny of a primary transformed or transfected cell, tissue, organ or whole plant that has been produced by any of the aforementioned methods, the only requirement being that progeny exhibit the same genotypic and/or phenotypic characteristic(s) as those produced in the parent by the methods according to the invention i.e. having modified growth characteristics. The invention accordingly also includes host cells having modified expression of a nucleic acid according to the invention and described herein before and/or having modified activity and/or level of a seedy1 protein or a protein according to the present invention, wherein said expression, level or activity is modified relative to corresponding wild type host cells. Preferred host cells according to the invention are plant cells or cells from insects, animals, yeast, fungi, algae or bacteria. The invention also extends to harvestable parts of a plant such as but not limited to seeds, flowers, stamen, leaves, petals, fruits, stem, stem cultures, rhizomes, roots, tubers and bulbs.

Preferably said host cells is transformed with a seedy1 encoding gene or a genetic construct as described above wherein the seedy1 nucleic acid is under the control of a seed-preferred promoter and more preferably the plants of the present invention carry an expression cassette comprising at least a part of seedy1 and at least a part of a constitutive promoter as mentioned hereinabove. The host cells, plants or the plant parts of the present invention can be identified by the presence of higher expression of a seedy1 gene and/or or a higher level and/or activity of a seedy1 protein. Further, particular plants of the present invention are recognizable by the presence of a seedy1 transgene.

The term "plant" as used herein encompasses whole plants, ancestors and progeny of the plants and plant parts, including seeds, shoots, stems, roots (including tubers), and plant cells, tissues and organs. The term "plant" also therefore encompasses suspension cultures, embryos, meristematic regions, callus tissue, leaves, gametophytes, sporophytes, pollen, and microspores. Plants that are particularly useful in the methods of the invention include all plants which belong to the superfamily *Viridiplantae*, in particular monocotyledonous and dicotyledonous plants including a fodder or forage legume, ornamental plant, food crop, tree, or shrub selected from the list comprising *Acacia spp., Acer spp., Actinidia spp., Aesculus spp., Agathis australis, Albizia amara, Alsophila tricolor, Andropogon spp., Arachis spp, Areca catechu, Astelia fragrans, Astragalus cicer, Baikiaea plurijuga, Betula spp., Brassica spp.,*

Bruguiera gymnorrhiza, Burkea africana, Butea frondosa, Cadaba farinosa, Calliandra spp, Camellia sinensis, Canna indica, Capsicum spp., Cassia spp., Centroema pubescens, Chaenomeles spp., Cinnamomum cassia, Coffea arabica, Colophospermum mopane, Coronillia varia, Cotoneaster serotina, Crataegus spp., Cucumis spp., Cupressus spp, Cyathea dealbata, Cydonia oblonga, Cryptomeria japonica, Cymbopogon spp., Cynthea dealbata, Cydonia oblonga, Dalbergia monetaria, Davallia divaricata, Desmodium spp, Dicksonia squarosa, Diheteropogon amplectens, Dioclea spp, Dolichos spp., Dorycnium rectum, Echinochloa pyramidalis, Ehrartia spp., Eleusine coracana, Eragrestis spp., Erythrina spp., Eucalyptus spp., Euclea schimperi, Eulalia villosa, Fagopyrum spp., Feijoa sellowiana, Fragaria spp., Flemingia spp, Freycinetia banksii, Geranium thunbergii, Ginkgo biloba, Glycine javanica, Gliricidia spp, Gossypium hirsutum, Grevillea spp., Guibourtia coleosperma, Hedysarum spp., Hemarthia altissima, Heteropogon contortus, Hordeum vulgare, Hyparrhenia rufa, Hypericum erectum, Hyperthelia dissoluta, Indigo incarnata, Iris spp., Leptarrhena pyrolifolia, Lespediza spp., Lettuca spp., Leucaena leucocephala, Loudetia simplex, Lotonus bainesii, Lotus spp., Macrotyloma axillare, Malus spp., Manihot esculenta, Medicago sativa, Metasequoia glyptostroboides, Musa sapientum, Nicotianum spp., Onobrychis spp, Ornithopus spp., Oryza spp., Peltophorum africanum, Pennisetum spp., Persea gratissima, Petunia spp., Phaseolus spp., Phoenix canariensis, Phormium cookianum, Photinia spp, Picea glauca, Pinus spp., Pisum sativum, Podocarpus totara, Pogonarthria fleckii, Pogonarthria squarrosa, Populus spp., Prosopis cineraria, Pseudotsuga menziesii, Pterolobium stellatum, Pyrus communis, Quercus spp., Rhaphiolepsis umbellata, Rhopalostylis sapida, Rhus natalensis, Ribes grossularia, Ribes spp., Robinia pseudoacacia, Rosa spp., Rubus spp., Salix spp., Schyzachyrium sanguineum, Sciadopitys verticillata, Sequoia sempervirens, Sequoiadendron giganteum, Sorghum bicolor, Spinacia spp, Sporobolus fimbriatus, Stiburus alopecuroides, Stylosanthos humilis, Tadehagi spp, Taxodium distichum, Themeda triandra, Trifolium spp., Triticum spp., Tsuga heterophylla, Vaccinium spp., Vicia spp., Vitis vinifera, Watsonia pyramidata, Zantedeschia aethiopica, Zea mays, amaranth, artichoke, asparagus, broccoli, Brussels sprouts, cabbage, canola, carrot, cauliflower, celery, collard greens, flax, kale, lentil, oilseed rape, okra, onion, potato, rice, soybean, straw, sugar beet, sugar cane, sunflower, tomato, squash tea, trees. Alternatively algae and other non-Viridiplantae can be used for the methods of the present invention. Preferably the plant according to the present invention is a crop plant selected from rice, maize, wheat, barley, millet, oats, rye, soybean, sunflower, canola, sugarcane, alfalfa, leguminosae (bean, pea), flax , lupinus, rapeseed, tobacco, tomato, potato, squash, papaya, poplar and cotton. Further preferably, the plant according to the present invention is a monocotyledonous plant, most preferably a cereal.

Advantageously, the present invention provides a method for modifying growth characteristics of a plant. According to further embodiments, the invention provides methods for modifying plant growth characteristics, wherein said modified growth characteristics are selected from any one or more of increased yield, increased biomass, modified plant architecture.

Further preferably, increased yield is increased seed yield en encompasses increased number of filled seeds and/or increased total seed weight per plant.
The term "increased yield" encompasses an increase in biomass in one or more harvestable parts of a plant relative to the total biomass of corresponding wild-type plants. The term also encompasses an increase in seed yield, which includes an increase in the biomass of the seed (seed weight) and/or an increase in the number of (filled) seeds and/or in the size of the seeds and/or an increase in seed volume, each relative to corresponding wild-type plants. An increase in seed size and/or volume may also influence the composition of seeds. An increase in seed yield could be due to an increase in the number and/or size of flowers. An increase in yield might also increase the harvest index, which is expressed as a ratio of the total biomass over the yield of harvestable parts, such as seeds.
The methods of the present invention are used to increase the seed yield of the plant and are therefore particularly favourable to be applied to crop plants, preferably seed crops and cereals. Therefore, the methods of the present invention are particularly useful for plants such as, rapeseed, sunflower, leguminosae (e.g. soybean, pea, bean) flax, lupinus, canola and cereals such as rice, maize, wheat, barley, millet, oats and rye.

Further preferably, increased biomass encompasses increased biomass of above-ground plant tissue, herein determined as above-ground plant area.
Additionally or alternatively, the plants according to the invention have more above ground area. Therefore, the methods of the present invention are additionally and/or alternatively particularly favourable to crops grown for the green tissue and/or grown for the above ground biomass. The methods of the present invention are particularly useful for increasing leaf size and number of grasses and forage crops (such as forage maize, clover, medicago alfalfa etc.). The methods of the present invention are also particularly useful for increasing the stem size of trees (for paper and pulp industry) and sugar cane.

Further preferably, said modified plant architecture encompasses increased number of panicles.
The methods of the present invention clearly change the appearance or morphology of a plant, including any one or more structural features or combination of structural features thereof. Therefore the plants according to the present invention have changed architecture when

compared to the wild-type plants. Other structural features, which may be altered by tl methods of the present invention include shape, size, number, position, texture, arrangemer and pattern of any cell, tissue or organ or groups of cells, tissues or organs of a plant, includir the root, leaf, shoot, stem or tiller, petiole, trichome, flower, inflorescence (f monocotyledonous and dicotyledonous plants), panicles, petal, stigma, style, stamen, polle ovule, seed, embryo, endosperm, seed coat, aleurone, fibre, cambium, wood, heartwoo parenchyma, aerenchyma, sieve elements, phloem or vascular tissue, amongst others.

The present invention also relates to use of a seedy1 nucleic acid and and/or protein i modifying growth characteristics and to compositions therefore. According to a particul embodiment, the methods of the present invention are used to change growth characteristi of a plant, wherein said growth characteristics is not disease or pathogen resistance, or is n stress tolerance.

The present invention also relates to use of a seedy1 nucleic acid and and/or protein as growth regulator, such as a herbicide or growth stimulator and to compositions therefore.
The present invention also relates to use of a seedy1 nucleic acid and and/or protein as target for an agrochemical compound, such as a herbicide or growth stimulator and t compositions therefore.

According to another aspect of the present invention, advantage may be taken of th nucleotide sequence capable of modifying expression of a seedy1 nucleic acid in breedin programmes. The nucleic acid may be on a chromosome, or a part thereof, comprising at lea the seedy1 nucleic acid and preferably also one or more related family members. In a example of such a breeding programme, a DNA marker is identified which may be genetical linked to a gene capable of modifying expression of a seedy1 nucleic acid in a plant, whic gene may be a gene encoding the seedy1 protein itself or any other gene which may directi or indirectly influence expression of the seedy1 gene and/or activity of the seedy1 protein itsel This DNA marker may then be used in breeding programs to select plants having change development.

Further the use of allelic variants as described herein-above are particular ly useful fc conventional breeding programmes, such as in marker-assisted breeding, which is als encompassed by the present invention. Such breeding programmes sometimes require th introduction of allelic variations in the plants by mutagenic treatment of a plant. One suitabl mutagenic method is EMS mutagenesis. Identification of allelic variants then may take plac by, for example, PCR. Tilling is preferred for identifying allelic variants. This is followed by

selection step for selection of superior allelic variants of the seedy1 sequence and which give rise to changed development in a plant. Selection, according to the method of the present invention, is typically carried out by monitoring development, differentiation and organ formation of plants containing different allelic variants of the seedy1 sequence, for example, different allelic variants of SEQ ID NO: 1 or of a seedy1 orthologue in that plant. Monitoring growth performance can be done in a greenhouse or in the field. Further optional steps include crossing plants, in which the superior allelic variant was identified, with another plant. This could be used, for example, to make a combination of interesting phenotypic features.

Therefore, mutations in the seedy1 gene may occur naturally, and may form the basis of the selection of plants showing accelerated rate of development, increased organ size and/or number, and/or early flowering.

Accordingly, as another aspect of the invention, there is provided a method for the selection of plants having modified growth characteristics, which method is based on the selection of better-performing allelic variants of the seedy1 sequence relative to the wild-type allele, and which give rise to modified growth characterstics in a plant.

According to a related embodiment of the invention, there is provided a method for breeding plants having modified growth characteristics, comprising the steps of:

a. Growing a plant with modified expression of a nucleic acid of the present invention as described herein above; and

b. Crossing the plant of (a) with another plant; and

c. selecting progeny of the cross of (b) having modified growth characteristic;

wherein said growth characteristics are modified relative to the corresponding wild-type plants.

The cross of step be is for example a cross with a commercial important germplasm.

Accordingly, the present invention also relates to the use of a seedy1 gene in breeding programmes.

Alternatively, the seedy1 gene itself can be used as a (genetic) marker to detect the presence or absence of a desired trait, or Quantitative Trait Locus (QTLs). In this application of the present invention the gene encoding seedy1 is genetically linked to the desired trait, and typically the phenotypes caused by the gene encoding a seedy1 are monitored in order to breed and select plants with the desired trait. This desired trait or QTL, may comprise a single gene or a cluster of linked genes that affect the desired trait.

In molecular biology it is standard practice to select upon transfection or transformation those individuals (or groups of individuals, such as bacterial or yeast colonies or phage plaques or

eukaryotic cell clones) that are effectively transfected or transformed with the desired genetic construct. Typically these selection procedures are based on the presence of a selectable or screenable marker in the transfected genetic construct, to distinguish the positive individuals easily from the negative individuals. Therefore, the seedy1 gene can also be used for these purposes, since introduction of this gene into a host cell results in changed development of said host cell.

The methods according to the present invention may also be practised by co-expression of a seedy1 gene in a plant with at least one other gene that cooperates with the seedy1 gene. Co-expression may be effected by cloning the genes under the control of a plant expressible promoter in a plant expressible vector and introducing the expression vector(s) into a plant cell using Agrobacterium-mediated plant transformation.

The methods according to the present invention result in plants having modified growth characteristics, as described hereinbefore. These advantageous characteristics may also be combined with other economically advantageous traits, such as further yield-enhancing traits, tolerance to various stresses, traits modifying various architectural features and/or biochemical and/or physiological features. Accordingly, the methods of the present invention can also be used in so-called "gene stacking" procedures.

Also the present invention encompasses a food product derived from any of the plants produced by the methods of the present invention. Further the invention also refers to the use of a product derived from any of the plants according to the present invention in animal feed and in food or in the production procedures thereof.

In a particular embodiment of the invention the plants with improved growth characteristics resulting in improved yield and/or biomass, are used to produce industrial enzymes and/ or pharmaceuticals. The production of such enzymes or pharmaceuticals in plants is aimed at high accumulation of the desired products in a particular and easy to harvest plant tissues, for example accumulation in the leaves and/or in the seeds. The plants of the present invention have more seeds and more above ground biomass, and therefore are capable of producing more industrial enzymes and/or pharmaceuticals in these tissues, more particularly in their green biomass and/or in their seeds. Accordingly, the present invention also provides a method for the production of enzymes and/or pharmaceuticals, which method comprises the modifying of expression of a seedy1 gene or the modifying of activity and/or level of a seedy1 protein. Further the invention relates to the use of plants according to the invention for the

production of industrial enzymes and pharmaceuticals and the invention extends to the industrial enzymes and pharmaceuticals produced according to these methods.

Description of the Figures

The present invention will now be described with reference to the following figures in which:

Figure 1 is a schematic presentation of the entry clone, containing CDS0689 within the AttL1 and AttL2 sites for Gateway® cloning in the pDONR201 backbone. CDS0689 is the internal code for the *Nicotiana tabacum* BY2 CDS0689 seedy1 coding sequence. This vector contains also a bacterial kanamycine-resistance cassette and a bacterial origin of replication.

Figure 2 is a map of the binary vector for the expression in *Oryza sativa* of the *Nicotiana tabacum* BY2 seedy1 gene (CDS0689) under the control of the rice prolamin RP6 promoter (PRO0090). This vector contains a T-DNA derived from the Ti Plasmid, limited by a left border (LB repeat, LB Ti C58) and a right border (RB repeat, RB Ti C58)). From the left border to the right border, this T-DNA contains: a selectable marker cassette for antibiotic selection of transformed plants; a screenable marker cassette for visual screening of transformed plants; the PRO0090 - CDS0689 -zein and rbcS-deltaGA double terminator cassette for expression of the *Nicotiana tabacum* BY2 seedy1 gene (CDS0689). This vector also contains an origin of replication from pBR322 for bacterial replication and a selectable marker (Spe/SmeR) for bacterial selection with spectinomycin and streptomycin.

Figure 3 is the alignment of seedy1 nucleic acids and EST's from different plant species. This alignment was made with the program align X of the NVTi software package. The motifs 1, 2, 3 and 4 are indicated with a line.

Figure 4 is the representation of nucleic acids, protein and motif sequences according to the invention.

Examples

The present invention will now be described with reference to the following examples, which are by way of illustration alone.

Unless otherwise stated, recombinant DNA techniques were performed according to standard protocols described in Sambrook (2001) Molecular Cloning: a laboratory manual, 3rd Edition Cold Spring Harbor Laboratory Press, CSH, New York; or in Volumes 1 and 2 of Ausubel *et al.*

(1988), Current Protocols in Molecular Biology, Current Protocols. Standard materials ai methods for plant molecular work are described in Plant Molecular Biology Labfase (1993) l R.D.D. Croy, published by BIOS Scientific Publications Ltd (UK) and Blackwell Scienti Publications (UK).

Example 1: cloning of the seedy1 encoding gene

A cDNA-AFLP experiment was performed on a synchronized tobacco BY2 cell cultu (*Nicotiana tabacum* L. cv. Bright Yellow-2), and BY2 expressed sequence tags that were c cycle modulated were identified and elected for further cloning. Subsequently, the expresse sequence tags were used to screen a tobacco cDNA library and to isolate the full-length cDN of interest, namely the cDNA coding for the seedy1 protein of the present inventic (CDS0689).

Synchronization of BY2 cells.

Tobacco BY2 (*Nicotiana tabacum* L. cv. Bright Yellow - 2) cultured cell suspension wa synchronized by blocking cells in early S-phase with aphidicolin as follows. Cultured ce suspension of *Nicotiana tabacum* L. cv. Bright Yellow 2 were maintained as described (Naga et al. Int. Rev. Cytol. 132, 1-30, 1992). For synchronization, a 7-day-old stationary culture wa diluted 10-fold in fresh medium supplemented with aphidicolin (Sigma-Aldrich, St. Louis, MO; mg/l), a DNA-polymerase a inhibiting drug. After 24 h, cells were released from the block t several washings with fresh medium and resumed their cell cycle progression.

RNA extraction and cDNA synthesis.

Total RNA was prepared by using LiCl precipitation (Sambrook et al, 2001) and poly(A+) RN was extracted from 500 mg of total RNA using Oligotex columns (Qiagen, Hilden, German according to the manufacturer's instructions. Starting from 1 mg of poly(A+) RNA, first-strar cDNA was synthesized by reverse transcription with a biotinylated oligo-dT25 primer (Gense Paris, France) and Superscript II (Life Technologies, Gaithersburg, MD). Second-strar synthesis was done by strand displacement with *Escherichia coli* ligase (Life Technologies and DNA polymerase I (USB, Cleveland, OH) and RNAse-H (USB).

cDNA-AFLP analysis.

Five hundred ng of double-stranded cDNA was used for AFLP analysis as described (Vos al., Nucleic Acids Res. 23 (21) 4407-4414, 1995; Bachem et al., Plant J. 9 (5) 745-53, 1996 The restriction enzymes used were BstYI and MseI (Biolabs) and the digestion was done i two separate steps. After the first restriction digest with one of the enzymes, the 3' en fragments were collected on Dyna beads (Dynal, Oslo, Norway) by means of their biotinylate

tail, while the other fragments were washed away. After digestion with the second enzyme, the released restriction fragments were collected and used as templates in the subsequent AFLP steps. For preamplifications, an MseI primer without selective nucleotides was combined with a BstYI primer containing either a T or a C as 3' most nucleotide. PCR conditions were as described (Vos et al., 1995). The obtained amplification mixtures were diluted 600-fold and 5 ml was used for selective amplifications using a P33-labeled BstYI primer and the Amplitaq-Gold polymerase (Roche Diagnostics, Brussels, Belgium). Amplification products were separated on 5% polyacrylamide gels using the Sequigel system (Biorad). Dried gels were exposed to Kodak Biomax films as well as scanned in a phosphoImager (Amersham Pharmacia Biotech, Little Chalfont, UK).

Characterization of AFLP fragments.

Bands corresponding to differentially expressed transcripts, among which the (partial) transcript corresponding to CDS0689, were isolated from the gel and eluted DNA was reamplified under the same conditions as for selective amplification. Sequence information was obtained either by direct sequencing of the reamplified polymerase chain reaction product with the selective BstYI primer or after cloning the fragments in pGEM-T easy (Promega, Madison, WI) or sequencing of individual clones. The obtained sequences were compared against nucleotide and protein sequences present in the publicly available databases by BLAST sequence alignments (Altschul et al., Nucleic Acids Res. 25 (17) 3389-3402 1997). When available, tag sequences were replaced with longer EST or isolated cDNA sequences to increase the chance of finding significant homology. The physical cDNA clone corresponding to CDS0689 was subsequently amplified from a commercial Tobacco cDNA library as follows.

Cloning of a tobacco CDS0689 seedy1 gene (CDS0689)

A c-DNA library with average inserts of 1,400 bp was made with poly(A+) isolated from actively dividing, non-synchronized BY2 tobacco cells. These library-inserts were cloned in the vector pCMVSPORT6.0, comprising an attB gateway cassette (Life Technologies). From this library 46,000 clones were selected, arrayed in 384-well microtiter plates, and subsequently spotted in duplicate on nylon filters. The arrayed clones were screened by using pools of several hundreds of radioactively labelled tags as probe (among which the BY2-tag corresponding to the sequence CDS0689). Positive clones were isolated (among which the clone reacting with the BY2-tag corresponding to the sequence CDS0689), sequenced, and aligned with the tag sequence. Alternatively, when the hybridization with the tag would fail, the full-length cDNA corresponding to the tag was selected by PCR amplification as follows. Tag-specific primers was designed using primer3 program (http://www-genome.wi.mit.edu/genome_software/other/primer3.html) and used in combination with the

common vector primer to amplify partial cDNA inserts. Pools of DNA from 50.000, 100.00 150.000, and 300.000 cDNA clones were used as templates in the PCR amplification Amplification product were isolated from agarose gels, cloned, sequenced and aligned wi tags. The vector comprising the sequence CDS0689 and obtained as described above, wa referred to as entry clone.

Example 2: Vector construction for transformation with PRO0090-CDS068 cassette

The entry clone was subsequently used in an LR reaction with p0830, a destination vect used for *Oryza sativa* transformation. This vector contains as functional elements within the DNA borders: a plant selectable marker; a plant screenable marker; and a Gateway cassett intended for LR in vivo recombination with the sequence of interest already cloned in the ent clone. The rice prolamin RP6 promoter for endosperm-specific expression (PRO0090) located upstream of this Gateway cassette.

After the LR recombination step, the resulting expression vector as shown in Fig. 2 can b transformed into *Agrobacterium* and subsequently into *Oryza sativa* plants. Transformed ric plants were allowed to grow and then examined for various parameters as described i example 3.

Example 3: Evaluation of T0, T1 and T2 transgenic rice plants transformed wit prolamin::seedy1 (PRO0090-CDS0689) and results

Approximately 15 to 20 independent T0 rice transformants were generated. The primar transformants were transferred from tissue culture chambers to a greenhouse for growing an harvest of T1 seed. 4 events, of which the T1 progeny segregated 3:1 for presence/absence c the transgene, were retained. For each of these events, approximately 10 T1 seedling containing the transgene (hetero- and homo-zygotes), and approximately 10 T1 seedling lacking the transgene (nullizygotes), were selected by monitoring screenable marke expression.

2 events with improved agronomical parameters in T1 were chosen for re-evaluation in T2 Seed batches from the positive plants (both hetero- and homozygotes) in T1, were screene by monitoring marker expression. For each chosen event, the heterozygote seed batches wer then selected for T2 evaluation. An equal number of positives and negatives within each see batch were transplanted for evaluation in the greenhouse. The total number of 120 seedy transformed plants were evaluated in the T2 generation . More particularly, two seedy transformed events have been selected, 60 plants per event of which 30 positives for th transgene, and 30 negative.

T1 and T2 plants were transferred to the greenhouse and evaluated for vegetative growth parameters and seed parameters, as described hereunder.

Table 1: overview of plants involved in T1 and T2 plant evaluation

Number and types of lines per evaluation	Number of plants per evaluation	Number of positive s per line	Number of negatives per line
4 lines of T1 plants	80	10	10
2 lines of T2 plants	120	30	30
1 line of T1 plants	20	10	10
1 line of T2 plants	60	30	30

Statistical analysis: t-test and F-test

A two factor ANOVA (analysis of variants) was used as statistical model for the overall evaluation of plant phenotypic characteristics. An F-test was carried out on all the parameters measured, for all of the plants of all of the events transformed with the gene of interest. The F-test was carried out to check for an effect of the gene over all the transformation events and to determine the overall effect of the gene or "global gene effect". Significant data, as determined by the value of the F-test, indicates a "gene" effect, meaning that the phenotype observed is caused by more than the presence or position of the gene. In the case of the F-test, the threshold for significance for a global gene effect is set at a 5% probability level.

To check for an effect of the gene within an event, i.e., for a line-specific effect, a t-test was performed within each event using data sets from the transgenic plants and the corresponding null plants. "Null plants" or "Null segregants" are the plants treated in the same way as the transgenic plant, but from which the transgene has segregated. Null plants can also be described as the homozygous negative transformants. The threshold for significance for the t-test is set at 10% probability level. Within one population of transformation events, some events can be under or above this t-test threshold. This is based on the hypothesis that a gene might only have an effect in certain positions in the genome, and that the occurrence of this position-dependent effect is not uncommon. This kind of gene effect may also be referred to as a "line effect of a gene". The p-value is obtained by comparing the t-value to the t-distribution or alternatively, by comparing the F-value to the F-distribution. The p-value stands for the probability that the null hypothesis (null hypothesis being "there is no effect of the transgene") is correct.

Vegetative growth measurements

The selected transgenic plants were grown in a greenhouse. Each plant received a uniqu barcode label to link unambiguously the phenotyping data to the corresponding plant. Th selected transgenic plants were grown on soil in 10 cm diameter pots under the followin environmental settings: photoperiod= 11.5 h, daylight intensity= 30,000 lux or more, daytim temperature= 28°C or higher, night time temperature= 22°C, relative humidity= 60-70% Transgenic plants and the corresponding nullizygotes were grown side-by-side at randor positions. From the stage of sowing until the stage of maturity each plant was passed severa times through a digital imaging cabinet and imaged. At each time point digital image (2048x1536 pixels, 16 million colours) were taken of each plant from at least 6 different angles The parameters described below were derived in an automated way from all the digital image of all the plants, using image analysis software.

(a) Aboveground plant area

Plant aboveground area was determined by counting the total number of pixels fron aboveground plant parts discriminated from the background. This value was averaged for th pictures taken on the same time point from the different angles and was converted to a physical surface value expressed in square mm by calibration. Experiments show that th aboveground plant area measured this way correlates with the biomass of plant parts abov ground.

b) Number of primary panicles

The tallest panicle and all the panicles that overlap with the tallest panicles when aligne vertically were counted manually, and considered as primary panicles.

Seed-related parameter measurements

The mature primary panicles of T1 and T2 plants were harvested, bagged, barcode-labelle and then dried for three days in the oven at 37°C. The panicles were then threshed and all th seeds were collected and counted. The filled husks were separated from the empty one using an air-blowing device. The empty husks were discarded and the remaining fraction wa counted again. The filled husks were weighed on an analytical balance. This procedur resulted in the set of seed-related parameters described below.

(c) Number of filled seeds

The number of filled seeds was determined by counting the number of filled husks tha remained after the separation step.

(d) Total seed yield per plant

The total seed yield was measured by weighing all filled husks harvested from a plant.

The results are given in % of difference between the positive plants and the corresponding nullizygotes (negative) plants of a transgenic line. The values given in table 2 to 4 are the average of 4 T1 lines and the average for T2 lines.

Table 2: overview of phenotypic data of seedy1 transgenic T1 and T2 plants for above ground area

	% difference between pos. and neg. plants for above ground area	
	T1 plants	T2 plants
4 lines	+ 12.5 %	
2 lines		+ 25.5 %

Table 3: overview of phenotypic data of seedy1 transgenic T1 and T2 plants for number of first panicles

	% difference between pos. and neg. plants for nr. of first panicles	
	T1 plants	T2 plants
4 lines	+ 32.25 %	
2 lines		+ 26.5 %

Table 4: overview of phenotypic data of seedy1 transgenic T1 and T2 plants for number of filled seeds

	% difference between pos. and neg. plants for nr. of filled seeds	
	T1 plants	T2 plants
4 lines	+ 59 %	
2 lines		+ 36.5 %

Table 5: overview of phenotypic data of seedy1 transgenic T1 and T2 plants for total seed weight per plant

	% difference between pos. and neg. plants for total seed weight per plant	
	T1 plants	T2 plants
4 lines	+ 70 %	
2 lines		+ 47 %

Example 4: Use of the invention in corn

The invention described herein is also be used in maize. To this aim, a seedy1 encoding gen for example a maize or other ortholog, is cloned under control of a seed-specific promote operable in maize, in a plant transformation vector suitable for Agrobacterium-mediated co transformation. Methods to use for corn transformation have been described in literatu (Ishida et al., Nat Biotechnol. 1996 Jun;14(6):745-50; Frame et al., Plant Physiol. 200 May;129(1):13-22).

Transgenic plants made by these methods are grown in the greenhouse for T1 see production. Inheritability and copy number of the transgene are checked by quantitative rea time PCR and Southern blot analysis and expression levels of the transgene are determine by reverse PCR and Northern analysis. Transgenic lines with single copy insertions of th transgene and with varying levels of transgene expression are selected for T2 see production.

Progeny seeds are germinated and grown in the greenhouse in conditions well adapted fc maize (16:8 photoperiod, 26-28°C daytime temperature and 22-24°C nighttime temperature as well under water-deficient, nitrogen-deficient, and excess NaCl conditions. Null segregant from the same parental line, as well as wild type plants of the same cultivar are used a controls. The progeny plants resulting from the selfing or the crosses are evaluated on differer biomass and developmental parameters, including, stem size, number of leaves, total abov ground area, leaf greenness, time to maturity, flowering time, time to flower, ear numbe harvesting time. The seeds of these lines are also checked on various parameters, such a grain size, total grain yield per plant, and grain quality (starch content, protein content and o content).

Lines that are most significantly improved compared to corresponding control lines ar selected for further field-testing and marker-assisted breeding, with the objective of transferrin the field-validated transgenic traits into commercial germplasm. The testing of maize fo growth and yield-related parameters in the field is conducted using well-established protocols Similarly, introgressing specific loci (such as transgene containing loci) from one germplasn into another is also conducted using well-established protocols.

Claims

1. An isolated nucleic acid encoding at least part of a seedy1 protein, wherein sa seedy1 protein comprises in the following order from N-terminus to C-terminus at least tw motifs selected from:
   (i) a motif having at least 80% sequence identity to the sequence represented by SE ID NO 15; and/or
   (ii) a motif having at least 80% sequence identity to the sequence represented by SE ID NO 16, and/or
   (iii) a motif having at least 80% sequence identity to the sequence represented by SE ID NO 17, a coiled coil motif; and/or
   (iv) a motif having at least 80% sequence identity to the sequence represented by SE ID NO 18.

2. An isolated nucleic acid selected from:
   (i) a nucleic acid represented by any of SEQ ID NO: 1, 5, 9, or the complement stran thereof;
   (ii) a nucleic acid encoding an amino acid sequence represented by SEQ ID NO: 2, ‹ 6, 8 or 10 or a homologue, derivative or active fragment of any of th aforementioned sequences;
   (iii) a nucleic acid capable of hybridising with a nucleic acid of (i) or (ii) above, whic hybridising sequence preferably encodes a protein having seedy1 protein activity;
   (iv) a nucleic acid which is degenerate from any one of the nucleic acids of (i) to (ii above as a results of the genetic code;
   (v) a nucleic acid which is an allelic variant of any one of the nucleic acids of (i) to (iv);
   (vi) a nucleic acid which is an alternative splice variant of any one of the nucleic acid of (i) to (v);
   (vii) a nucleic acid encoding a protein which has at least 21%, 22%, 23%, 24%, 25% 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39% 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 60%, 70%, 75% 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% sequence identity to any one c more of the sequences defined in (i) to (vi), which protein preferably encodes protein having seedy1 activity;
   (viii) a portion of a nucleic acid according to any of (i) to (vii) above, which portio preferably encodes a protein having seedy1 activity.

3. An isolated seedy1 protein comprising in the following order from N-terminus to C-terminus at least two motifs selected from:
   (i) a motif having at least 80% sequence identity to the sequence represented by SEQ ID NO 15; and/or
   (ii) a motif having at least 80% sequence identity to the sequence represented by SEQ ID NO 16, and/or
   (iii) a motif having at least 80% sequence identity to the sequence represented by SEQ ID NO 17, a coiled coil motif; and/or
   (iv) a motif having at least 80% sequence identity to the sequence represented by SEQ ID NO 18.

4. An isolated seedy1 protein, comprising
   a. a polypeptide with an amino acid sequence as presented in any one of SEQ ID NO 2, 4, 6, 8 or 10;
   b. a polypeptide with an amino acid sequence which has at least 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% sequence identity to any one or more of the amino acid sequence as described in (a)
   c. a polypeptide which is a homologue, derivative, immunologically active and/or functional fragment of a protein as defined in any of (a) or (b).

5. An isolated genetic construct comprising:
   (i) a nucleic acid of claim 1 or 2 or a nucleic acid encoding a protein according to claim 3 or 4;
   (ii) one or more control sequence capable of regulating expression of the nucleic acid of (i); and optionally
   (iii) a transcription termination sequence.

6. An isolated genetic construct according to claim 5, wherein said nucleic acid is represented by SEQ ID NO 1 or a portion thereof or by sequences capable of hybridising therewith, which nucleic acid is preferably from a dicotyledonous plant, further preferably from the family *Solanaceae*, more preferably the nucleic acid is from *Nicotiana*, or is a nucleic acid encoding an amino acid sequence represented by SEQ ID NO: 2 or a homologue thereof having at least, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%

45%, 46%, 47%, 48%, 49%, 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%; 98' or 99% sequence identity with SEQ ID NO 2, or derivative or active fragment thereof.

7. An isolated genetic construct according to claim 5 or 6, wherein said control sequenc is a plant promoter, preferably a seed-specific promoter, more preferably an endosperr specific promoter, more preferably a promoter isolated from a gene encoding a seec storage protein, most preferably a promoter isolated from a prolamin gene, such as fc example the rice prolamin promoter of SEQ ID NO 14.

8. An isolated genetic construct according to claim 7, comprising an nucleic acid selecte from:

a. A nucleic acid having at least a part of a seed-preferred promoter and the nuclei acid as defined in claims 1 or 2; and
b. A nucleic acid as presented in SEQ ID NO 13, or the complementary strand thereo
c. a nucleic acid which is degenerated as a result of the genetic code from any of th nucleic acids of (a) or (b);
d. a nucleic acids which is an allelic variant of any of the nucleic acids of (a) or (b);
e. a nucleic acid which hybridises to any of the nucleic acids of (a) or (b).

9. Method for modifying growth characteristics of a plant, comprising modifying in a plar expression of a nucleic acid according to claim 1 or 2 and/or modifying in a plant lev‹ and/or activity of a seedy1 protein according to claim 3 or 4, wherein said growt characteristics are modified relative to corresponding wild-type plants.

10. Method according to claim 9, wherein said modification is effected by recombinar means and/or chemical means.

11. Method according to claim 9 or 10, comprising introducing, into a plant, a nucleic aci capable of modifying expression of a gene encoding a seedy1 protein and/or capable c modifying levels and/or activity of a seedy1 protein.

12. Method according to claim 11, wherein said nucleic acid is a nucleic acid of claim 1 c 2.

13. Method according to claim 12, wherein said nucleic acid is as represented by SEQ II NO 1 or a portion thereof or sequences capable of hybridising therewith, which nucleic aci is preferably from a dicotyledonous plant, further preferably from the family *Solanaceae*

more preferably the nucleic acid is from *Nicotiana*, or is a nucleic acid encoding an amino acid sequence represented by SEQ ID NO 2 or a homologue thereof having at least, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% sequence identity with SEQ ID NO 2, or derivative or active fragment thereof.

14. Method according to any of claims 9 to 13, wherein said nucleic acid is an allelic variant of a nucleic acid encoding the seedy1 protein of claim 2 or 3, or wherein said seedy1 protein is encoded by an allelic variant.

15. Method according to any of claims 9 to 13, wherein said nucleic acid is an alternative splice variant of a nucleic acid encoding the seedy1 protein of claim 2 or 3, or wherein said seedy1 protein is encoded by a splice variant.

16. Method according to any claims 9 to 13, wherein said nucleic acid is comprised on at least a part of an artificial chromosome, which artificial chromosome preferably also comprises one or more related gene family members.

17. A method for the production of a transgenic plant having modified growth characteristics, which method comprises:
   (i) Introducing into a plant cell a nucleic acid of claim 1 or 2;
   (ii) cultivating the plant cell under conditions promoting regeneration and mature plant growth.

18. Method according to claim 17, wherein said nucleic acid is overexpressed in a plant.

19. Method according to any of claims 17 or 18, wherein said nucleic acid is introduced in the sense direction into a plant.

20. Method according to any of claims 17 to 19, wherein expression of said nucleic acid is driven by a seed-preferred promoter.

21. Method according to any of claims 17 to 20, wherein expression of said nucleic acid is driven by a promoter of a seed-storage protein.

22. Method according to claim 21, wherein the promoter is a prolamin promoter.

23. Method for breeding plants having modified growth characteristics, comprising th steps of:
    (i) Growing a plant with modified expression of a nucleic acid of claim 1 or 2; and
    (ii) Crossing the plant of (i) with another plant; and
    (iii) selecting progeny of the cross of (ii) having modified growth characteristics;

wherein said growth characteristics are modified relative to the corresponding wild-typ plants.

24. Method according to any of claims 9 to 23, wherein said modified growth characteristic is selected from any one or more of increased yield, increased biomass, modified plar architecture, each relative to corresponding wild type plants.

25. Plants obtainable by a method according to any of claims 9 to 24, which plants hav modified growth characteristics.

26. A host cell having modified expression of a nucleic acid of claim 1 or 2 and/or havin modified activity and/or level of a protein of claim 2 or 4, wherein said expression, level c activity is modified relative to corresponding wild type host cells.

27. A host cell according to claim 26, comprising a nucleic acid of claim 1 or 2 or and/c containing a genetic construct of any one of claims 5 to 8.

28. A host cell according to claims 26 or 27, which is a cell from a plant, insect, animal yeast, fungus, algae or bacterium.

29. A plant having modified growth characteristics, comprising the host cell of claim 28.

30. A plant according to claim 25 or 29, wherein said plant is a crop plant selected fron rice, maize, wheat, barley, millet, oats, rye, soybean, sunflower, canola, sugarcane, alfalfa leguminosae (bean, pea), flax , lupinus, rapeseed, tobacco, tomato, potato, squash papaya, poplar and cotton, preferably said plant is a monocotyledonous plant, mor preferably a cereal.

31. Plant part, preferably a harvestable plant part, a propagule or progeny from a plan according to claim 25, 29 or 30.

32. Use of a nucleic acid of claim 1 or 2 and/or of a seedy1 protein of claim 3 or 4 in modifying the growth characteristics of a plant.

33. Use of a nucleic acid of claim 1 or 2 or of a seedy1 protein of claim 3 or 4 as

(i) a growth regulator, such as a herbicide or a growth stimulator; or
(ii) as a target for an agrochemical compound, such as a herbicide or a growth stimulator; or
(iii) use of a nucleic acid of claim 1 or 2 in breeding programs; or
(iv) use of a plant according to any one of claims 25, 29 or 30, or use of a product derived therefrom in animal feed or in a food product.
(v) Use of plants according to any one of claims 21, 29 or 30 for the production of (industrial) enzymes, pharmaceuticals or plant effective agents.

34. (Industrial) enzymes, pharmaceuticals and plant effective agents, produced in a plant according to any one of claims 25, 29 or 30.

35. A food product derived from a plant according to any one of claims 21, 29 or 30.

36. A growth regulating composition, such as a herbicide or a growth stimulator, comprising a nucleic acid of claim 1 or 2 and/or comprising a seedy1 protein of claim 3 or 4, said composition further comprising substances normally added to growth regulating compositions.

Abstract

Plants having modified growth characteristics and a method for making the same

The present invention concerns a method for modifying growth characteristics of a plant, in particular increasing biomass and seed yield and modifying plant architecture, by modifying expression of a seedy1 nucleic acid and/or modifying levels and/or activity in a plant of a seedy1 protein. The invention also relates to isolated nucleic acids encoding a seedy1 protein, which seedy1 protein has a coiled coil domain as represented in SEQ ID NO 17 and three conserved motifs as presented in SEQ ID NO 15, 16 and 18. Further the invention relates to transgenic plants having modified growth characteristics, which plants have modified expression of a nucleic acid encoding a seedy1 protein or modified levels and/or activities of a seedy1 protein.

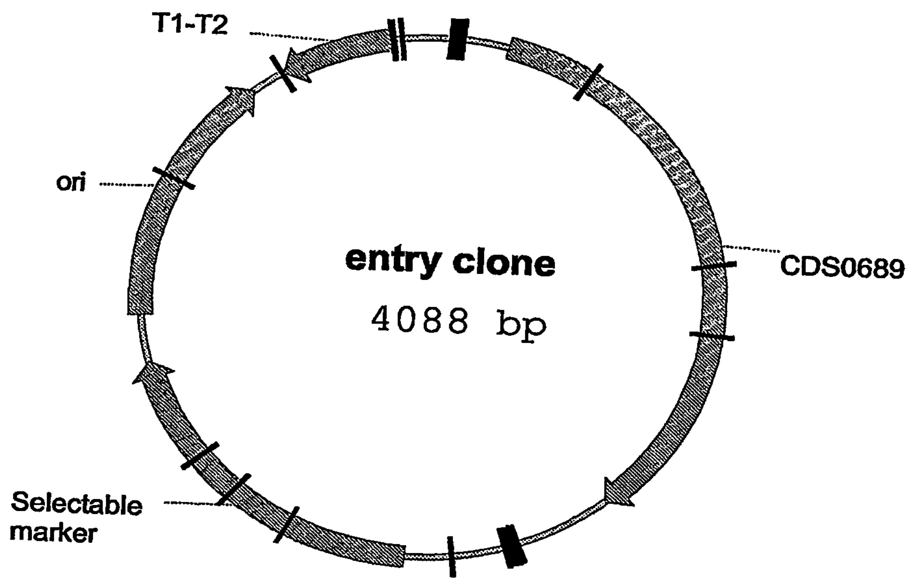

FIGURE 1

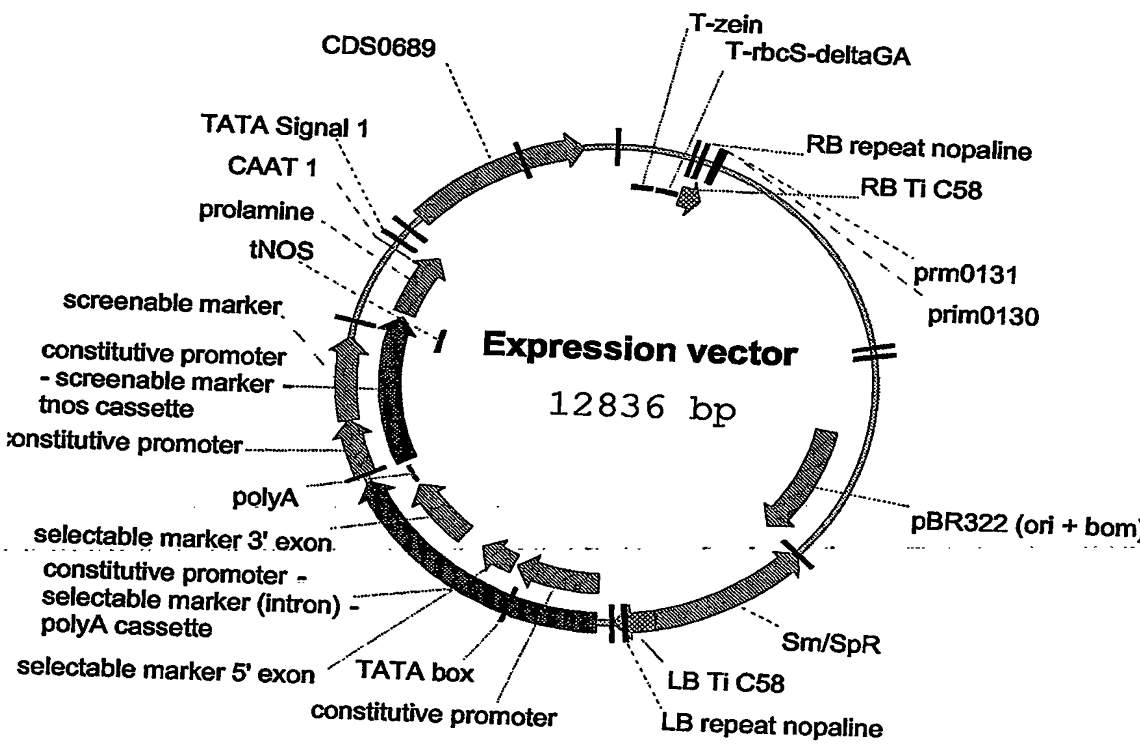

FIGURE 2

```
                                              1            Motif 1             40
                    CDS0689  (1)   ----MSSLQYPEGIDPAEVQIWNNAAFDNGDSEDLSSIKR
                 CDS0689 At  (1)   ---MTSIEATETLNAPPKEQIWNNAAFIDGDSQIISAIEA
  CDS0689 Medicago trunculata (1)  ---MNNTNNNNILLHSTQVQVWNNAAFDGEDFAMNSESDS
                 CDS0689 Os  (1)   ---------MEEDPLIPLVHVWNNAAFIDSSCSRSAWIPQ
         CDS0689 Ta variant  (1)   --------MMEEDPLIPLVHVWTIAAFISSSSSSAWHAHA
                 CDS0689 So  (1)   ---------MEEDPLIPLVHVWNNAAFIHASSSAWHAHSP
       CDS0689 Zm variantrev (1)   ---------MEEDPLIQLVHVWSIAACINAASSSVCHAH
         CDS0689 So variant  (1)   ---------MEEDPLIPLVHVWNNAPFIHASYSAWHAHSP
            CDS0689 Bn est1  (1)   ----MTSTEHTETLNAPEEQIWNNAAFIDGDSNLISAIEA
CDS0689 Eschscholzia californica (1) ------MLEISETLNEPDEQIWNIAAFDSGSIDNHTTAIK
            CDS0689 Ga est1  (1)   ----MSSLQYPESFNVPEEQVWNNAAFDNGDSEDINAIKD
                 CDS0689 Pt  (1)   ---MSSSLQYPEVVDAPEVQIWNNAAFDNGISEGSLNIKS
  CDS0689 Plumbao zeylanica (1)   MNEVLHIQEAARTDSSTDHQIWNNAAFISGESEDSPVVID
    CDS0689 Citrus sinensis (1)   ----MSSLQYPEPTLNGQEEQIWNNAAFDNGDSEDSTAMKG
                  Consensus  (1)       I   E   IPEVQIWNNAAFD GDS   S AI

                                             41                  Motif 2       80
                    CDS0689  (37)  --SWSPLKPLSVRPSDSFESDLSSKENQTPLFENSSVNLS
                 CDS0689 At  (38)  SSWSHLN-------ESFDSECS--KENQFPISVSSSLQSS
  CDS0689 Medicago trunculata (38) -----------------------IKENLNPSAFN------
                 CDS0689 Os  (32)  ---SP-----------AVAAVRKGDKENHRPEVVD------
         CDS0689 Ta variant  (33)  -----------------TPVRRGEGKRRPAETN------
                 CDS0689 So  (32)  ------------VPASARRAEGDKENHRPDPDP------
       CDS0689 Zm variantrev (32)  -----------SPAPASARPGEGDKEMLRPEPDV------
         CDS0689 So variant  (32)  -------------ARASAGHEAEGDKENHRPDPDP------
            CDS0689 Bn est1  (37)  ----------SWSNLNESFDSECSKENQIPVSVSSSLKSS
CDS0689 Eschscholzia californica (35) ASSSPLKPIVLNQSEPSILDSIYTKENQTPSCCISPVRTK
            CDS0689 Ga est1  (37)  SWCNFNS-----GSVNQSLESDGSKENQSPLWIKSPVSFK
                 CDS0689 Pt  (38)  ------------SWWNQSLESDASKENLSPVCEQSSPVFV
  CDS0689 Plumbao zeylanica (41)   --F------SAPNLSQSLLSDSSIKENLSPSLAEMPHPAK
    CDS0689 Citrus sinensis (37)  SWANLKS-----VYMNQSLESDCSKENLSPRLNKSPTSSL
                  Consensus  (41)                S  ESDG KEN  P    S

                                             81                                120
                    CDS0689  (75)  SPLPIKPLNPNGALENSRLKPNKPNSKQSLDEMAARKSGK
                 CDS0689 At  (69)  VSITEAPS--------------AKSKTVKTKSAADRSKK-
  CDS0689 Medicago trunculata (49) ---IVPS-----------------------------SNK-
                 CDS0689 Os  (53)  ---VAAG---------------------------------
         CDS0689 Ta variant  (50)  ----------------------------------------
                 CDS0689 So  (54)  ----------------------------------------
       CDS0689 Zm variantrev (55)  ----------------------------------------
         CDS0689 So variant  (54)  ----------------------------------------
            CDS0689 Bn est1  (67)  VSFSTDDP----------------IRCGKVK-----EKPH
CDS0689 Eschscholzia californica (75) SPLPIKPLHPNG----------------------------
            CDS0689 Ga est1  (72)  STASVVKP-------LSSKNVTGNTREPFSAKMKSGVCKE
                 CDS0689 Pt  (66)  NSSKPAKPLQ------------------------------
  CDS0689 Plumbao zeylanica (73)   SPMQK-----------------------------------
    CDS0689 Citrus sinensis (72)  KSCVPNKPLQVN----------SSVKNSQMKQLK--SVSK
                  Consensus  (81)

                                             121      Motif 3: coiled coil     160
                    CDS0689  (115) GNDFRDEKKIDEEIEEIQMEISRLSSRLEALRIEKAEKTV
                 CDS0689 At  (94)  -------RDIDAEIEEVEKEIGRLSTKLESLRIEKAEQTA
  CDS0689 Medicago trunculata (56) -------RTIDDEIAEIESEIKRLISKLELLRVEKAERKI
                 CDS0689 Os  (57)  -------YDVEAEIGHIEAEILRLSSRLHHLRVSKQPEPN
         CDS0689 Ta variant  (50)  --------DADAEIARIEAEILRLSSRLHHLRVSKGHDAK
                 CDS0689 So  (54)  --------DVEAEIGHIEAEILRLXSRLHHLRTSKQSEPS
       CDS0689 Zm variantrev (55)  ---------EEERHIEAEILRLSLRLHHLRTSQQLQPP
         CDS0689 So variant  (54)  --------DVEAEIGHIEAEILRLSSRLHHLRTSKQSEPP
            CDS0689 Bn est1  (86)  KTGKVRHGDIDAEIEEVEKEMNRLSIRLESLRIEKAEQIA
CDS0689 Eschscholzia californica (87) ----------------------------------------
            CDS0689 Ga est1  (105) EEKKRDEKKIDMEIEEIEKEVARLSAKLESLRIEKPNIMQ
                 CDS0689 Pt  (76)  ----------------------------------------
  CDS0689 Plumbao zeylanica (78)   ----------------------------------------
    CDS0689 Citrus sinensis (100) EEETRDERKIDIEIEEIEKEISRLSSRLEALRIEKIDIKT
                  Consensus  (121)           ID EI   IE EI RLSSRL  LRL K
```

FIGURE 3

```
                                      401                                    440
                   CDS0689      (339) IQSSVVRKRSLPENDKDESKRNDKKRSLSVGKTRVSQTES
                CDS0689 At      (294) TGEKDVRKRSLPEDEEKENHKRSEKR-------RASDESN
CDS0689 Medicago truncatula     (284) GGD--ARKRSFSEN-------------------------N
                CDS0689 Os      (297) AAAAATAKRMAGSSKMRVIPSRYSNTPGASLGSSGAQERR
       CDS0689 Hv contig 123    (111) ISTASTCRRPAGSSKVRVVPSRYSLMPGASLG-AATQDGR
            CDS0689 Ta contig     (1) --------------------GRYSLMPGASLGAASQERRR
           CDS0689 Zm partial    (87) TSNVATTKRPAGSSKVRVVPSRYSNPPGSSLAAVTQGNRC
           CDS0689 Sacc sp 3'    (84) TSNAATAKRPAGSSKVRVVPSRYSNTPGSYLAAVSQDKRS
        CDS0689 Pinus taeda 3'    (1) ----------------------------------------
                  Consensus     (401)     KR        R     RYSL PGA LG
                                      441                                    480
                   CDS0689      (379) KNL---GTESRVKKRWEIPSEIVVHGNTESEKSPLSNIVK
                CDS0689 At      (327) K------SEGRVKKRWEIPSEVDLYSSGEN--GDESPIVK
CDS0689 Medicago truncatula     (297) KGL---GSEIRAKKRWEIPIEEVDVSG--------FVM--
                CDS0689 Os      (337) RKQSLPGSSGDANQNEEIRAKVIEPSN--DPLSPQTISKV
       CDS0689 Hv contig 123    (150) RKESLPGSTGSTGQKEEIKAVPTEPVD--DDLSPESIDKV
            CDS0689 Ta contig    (21) KESLPGSSGGAGQKEEEIKAMPTEPVD--DDLSPESIDKV
           CDS0689 Zm partial   (127) KQS-----LPGSATETRVNLTEPPNDE----LSPEEIAKV
           CDS0689 Sacc sp 3'   (124) KQS-LPGPASAASQREEIRAKLTEPSK--DELSPETVAKV
        CDS0689 Pinus taeda 3'    (1) ------------XEARIVFGTGNSAIMAGGTKAPDTIERH
                  Consensus     (441) K     GS      R EI A    E         LSPESL KV
```

Motif 4

```
                                      481                                    520
                   CDS0689      (416) PDLLPRIRIARCVNEILRDSGPAKRMIELIGKKSFFSSDE
                CDS0689 At      (359) --EVPKIRTSRRVGGSPRDSGAAKRVAELQAK--------
CDS0689 Medicago truncatula     (324) ---LPRISTMRFVDESPRDSCAVKRVAELNGKRSVFCDED
                CDS0689 Os      (375) AEMLPKIRTMPPPDESPRDSGCAKRVAELVGKRSFFTAAA
       CDS0689 Hv contig 123    (188) AELLPRIRTMPRPNEIPRDSGCAKRAADLVGKRSFFAAAA
            CDS0689 Ta contig    (59) AELLPRIRTMPPPDEIPRDSGCAKRAADLVGKRSFFAAAA
           CDS0689 Zm partial   (158) AELLPRIRTMPPSDESPRDSGCAKRVADLVGKRSFFTAAG
           CDS0689 Sacc sp 3'   (161) AELLPRIKTMPASDESPRDSGCAKRVADLVGKRSFFTXAA
        CDS0689 Pinus taeda 3'   (29) KMKLPKIKTVRFTTESPRDSGCIKREIDRIGKKSFFAPDG
                  Consensus     (481) AELLPRIRTMP  DESPRDSGCAKRVADLVGKRSFF AAA
                                      521                                    560
                   CDS0689      (456) ---------DKEPPVCQVLSFAEEDAEEE-----------
                CDS0689 At      (389) ---------DRNFTFCQDLKFEE-----------------
CDS0689 Medicago truncatula     (361) EEERVMVEEEGGSVCQVLNFAEDEDDDDDYGEQG------
                CDS0689 Os      (415) -------------EDGEALSVEEPEAVAEA----------
       CDS0689 Hv contig 123    (228) AGD---GSAISSYQARVLEAEAPEEAAAAGALSDEAAAAG
            CDS0689 Ta contig    (99) AGD---CSAISSYQARVLEAEAPEEAAAAAAEALGDEAASA
           CDS0689 Zm partial   (198) DDG----NLVTPYQARVVELESPEAAAEEAEA--------
           CDS0689 Sacc sp 3'   (201) EDG----NFVTPYQAPVGEL--------------------
        CDS0689 Pinus taeda 3'   (69) ITST--PSIDXXDAGKPLRRESVHEIXHAXXX--------
                  Consensus     (521)                ARVLE EA E
                                      561                            592
                   CDS0689      (476) --------------------------------
                CDS0689 At      (403) --------------------------------
CDS0689 Medicago truncatula     (395) --------------------------------
                CDS0689 Os      (432) --------------------------------
       CDS0689 Hv contig 123    (265) ALSDEAAAAAAAAEALSDEAAAAEALSDEAAA
            CDS0689 Ta contig   (136) GEALGDEAAA----------------------
           CDS0689 Zm partial   (226) --------------------------------
           CDS0689 Sacc sp 3'   (217) --------------------------------
        CDS0689 Pinus taeda 3'   (99) --------------------------------
                  Consensus (561)
```

FIGURE 3 (continued)

SEQ ID NO 1: *Nicotiana tabacum* seedy1 coding sequence (CDS0689)

atgagtgtgttacaatacccagaagggattgacccagcagatgttcagatatggaacaatgc
agcatttgataatggagattctgaagatttgtcttcgctgaaacgttcttggtctcctctga
aacccctttcggttaggccatcagattcctttgaatctgatttgtcaagtaaggaaaatcaa
actcctttatttgagaattcatctgttaatctctcatctccgttacccataaagccacttaa
ccctaatggggctctggaaaattcaagactcaagccgaacaagcccaattccaaacagagtc
ttgatgagatggcggctagaaagagcggaaagggaaatgatttccgtgatgagaagaaaata
gacgaggaaattgaagaaattcagatggagattagtaggttgagttcaagattagaggcttt
gagaattgaaaaggctgagaaaactgttgctaagactgttgaaaagcgaggaagggttgtgg
cagcaaagtttatggagccaaaacaaagtgttattaagattgaagagcgtatatcaatgagt
gcaagaacaaaggtggagcagagaaggggtcttagtttaggaccatctgagatttttactgg
aacgcggcggcgagggttgagtatggggccatcagatattctagcagggacaacaaaggcac
ggcaattgggaaagcaagagatgattattactcctattcagccaatacaaaacaggcgaaag
tcgtgtttttggaagcttcaagagattgaagaagagggaaaaagttcaagccttagtcctaa
atcaagaaaaactgctgcaagaacaatggttacaacaaggcaggcagttactacaattgcat
caaagaagaatttgaaaaaagatgatggacttttgagttcagttcagccaaagaagttgttt
aaagatctcgaaaagtctgctgctgctaataagaagccccagaggccggggagggttgtggc
tagtaggtataatcagagtacaattcagtcatcagtagtgagaaagaggtctttacctgaaa
atgataaggatgagagtaagagaaatgataagaaacggtcgttatctgtagggaaaacgcgt
gtgtctcaaactgagagcaagaatttgggtactgaaagtagggtgaaaaagagatggaaat
tcctagtgagattgtagttcatggaaacacagagagtgagaaatctccactaagcattattg
tgaagcctgatttgcttccgcgaattaggattgctcggtgtgtgaatgagactcttagggat
tctggacctgctaaaagaatgatagagttgataggcaagaaatcgtttttcagtagtgatga
agataaggagccacctgtctgtcaagttttaagttttgcagaggaagatgctgaagaggaa**t**
aa

SEQ ID NO 2: *Nicotiana tabacum* seedy1 protein (CDS0689)

MSVLQYPEGIDPADVQIWNNAAFDNGDSEDLSSLKRSWSPLKPLSVRPSDSFESDLSSKENQ
TPLFENSSVNLSSPLPIKPLNPNGALENSRLKPNKPNSKQSLDEMAARKSGKGNDFRDEKKI
DEEIEEIQMEISRLSSRLEALRIEKAEKTVAKTVEKRGRVVAAKFMEPKQSVIKIEERISMS
ARTKVEQRRGLSLGPSEIFTGTRRRGLSMGPSDILAGTTKARQLGKQEMIITPIQPIQNRRK
SCFWKLQEIEEEGKSSSLSPKSRKTAARTMVTTRQAVTTIASKKNLKKDDGLLSSVQPKKLF
KDLEKSAAANKKPQRPGRVVASRYNQSTIQSSVVRKRSLPENDKDESKRNDKKRSLSVGKTR
VSQTESKNLGTESRVKKRWEIPSEIVVHGNTESEKSPLSIIVKPDLLPRIRIARCVNETLRD
SGPAKRMIELIGKKSFFSSDEDKEPPVCQVLSFAEEDAEEE

FIGURE 4

SEQ ID NO 3: *Oryza sativa* seedy1 coding sequence
atggaggaggacccgctcatcccgctggtccacgtctggaacaacgccgccttcgacgactc
ctcgtgttccagatcggcttggctccccaaagccccgccgtcgcggccgtccgcaagggcg
acaaggagaatcaccgccccgaggttgttgatgtcgccgccggctacgacgtcgaggccgag
atcggccacatcgaggcggagatcctgcgcctctcgtcccggctccaccatctccgcgtctc
caagcagccggagcccaaccgcgacgacgctccgatgggggagatggtcgcgaaggtgaggc
cccggccgaggggcctcagcctcgggcccctggatgtgatctccatcgtcaatcgtgagaag
catccgctgcgcaccaagcagcctccggcgacgcggggcaggggctcagcctcgggcccat
ggagatcgccgcggcgaaccctagggtgcccgcggcggcgcagcatcagcaacagcaacgcg
ctggcacggcgcggatcctgaagccaatcaaggagcctccggtgcagcgtcgcaggggcgtc
agcctcgggccgttggagatccaccacggcgtcggcagcaaggcaccagcggcggcgcgagc
caagccgttcaccaccaagctcaacgccattcgagaagaaacccgaccctccaagcaattcg
ccgtccccgccaagccatggccgtcgagcaatacaaggcagacactggactcgaggcaagga
acagcagcaagtcgagcgaaggcgaggagcccgagccccaggcccaggaggcaatccaatgg
caaggctactgacacaaggggaggcaacaaggtggtggatgagctcaagcccaaaggtgcgt
cgtcaagtcagagcggcagcgccgccgccgccgccactgccaagaggatggcggggagctcc
aagatgagggtcatcccgagccgctacagcctcactcctggcgcttcccttggaagcagtgg
agcacaggagaggcgacgcaagcagtctctcccaggatcatcaggggatgcgaaccagaatg
aggaaatcagagcgaaggtcatcgagccttccaatgatccactctctcctcaaacgatctcc
aaggttgctgaaatgctcccaaagatcaggaccatgccgcctcctgacgagagccctcgcga
ttccggatgcgccaagcgggttgccgaattggtcgggaagcgctcgttcttcacggctgcag
ccgaggacgggcgggcgctcgacgtcgaagcacccgaggcggtcgcagaagcttgagatgaa
ccaccatggtttgatccgttccttccatcagctc

SEQ ID NO 4: *Oryza sativa* seedy1 protein
MEEDPLIPLVHVWNNAAFDDSSCSRSAWLPQSPAVAAVRKGDKENHRPEVVDVAAGYDVEAE
IGHIEAEILRLSSRLHHLRVSKQPEPNRDDAPMGEMVAKVRPRPRGLSLGPLDVISIVNREK
HPLRTKQPPATRGRGLSLGPMEIAAANPRVPAAAQHQQQQRAGTARILKPIKEPPVQRRRGV
SLGPLEIHHGVGSKAPAAARAKPFTTKLNAIREETRPSKQFAVPAKPWPSSNTRQTLDSRQG
TAASRAKARSPSPRPRRQSNGKATDTRGGNKVVDELKPKGASSSQSGSAAAAATAKRMAGSS
KMRVIPSRYSLTPGASLGSSGAQERRRKQSLPGSSGDANQNEEIRAKVIEPSNDPLSPQTIS
KVAEMLPKIRTMPPPDESPRDSGCAKRVAELVGKRSFFTAAAEDGRALDVEAPEAVAEA

FIGURE 4 (continued)

SEQ ID NO 5: *Medicago trunculata* seedy1 encoding sequence
aaaaacgttaaggactaaaatataataaaatttaagtagggattcataatggaagcacccc
tatttacagggatcttaaatataattaaccctaatatttatgacagaaacccttttgaaatc
acatcggagcgtgtatgagtagccgtttcacatccaacggccagtaagagcgtaactttatt
tcttccctcttcaatctccaacggtcacataatctcttccaaatacaataattccctcttt
caacctcactcttcatttcttcaacccaaacccaaaaaactaatcagattcttcttaaatct
tgaaacctttctcccaaaagcacttaaataaaaaagcacttaaccatgaataacacaaacaa
caacaacattcttcttcattccacacaggttcaagtgtggaacaacgcagcattcgatggtg
aagatttcgccatgaattcatcttctgattccatcaaagagaatctaaacccatccgcattc
aacattgttccttcttcaaacaaaagaactattgatgatgaaattgcggaaattgaaagtga
aattaagcgattaacttcgaagctggaattgcttcgtgttgaaaaagctgaaagaaaaatcg
cttctgaaaagcgtgttagtggaattggtactggaagaatagtagcagcgaagtttatggaa
ccgaagaaaaacgttacaccgaaacgaaacggtgtcgttttcaaggaggagacaccgaaacg
aaacggtgtcgtttcggatacgccgaaatctagggttaattggagaagagggatgagtttag
gtccgatggagattgccgggaaagtgatggcaccgccggcgatgacgattactccggcgacg
gtgaatcggaggaagtcttgtttctggaaaccgcaggaaagttgtgaagtaatgccgtcggg
gattactccggcgacggtgaataggaggaaatcttgtttttgaaacctcaagaaagttgtg
aagaaaatcgaagaaaaacgatttgcaaaccgaatttgaatttgaattcaaattcagttaat
tctgcggttggatcgattaagcgtgtgaagaagaaagatgaagaaattgctcaggttcaacc
gaagaagctgtttgaaggtgaaaaatcagtgaagaaatcgttgaaacaaggtagaattgttg
caagccggtataattccggtggtggtggtggtgatgcgaggaaaagatcgttttcggagaat
aataagggtttagggagtgaaatcagggctaagaagagatgggagataccaattgaagaagt
ggatgtgagtggttttgttatgttaccgaagatttcgacaatgaggtttgttgatgagagtc
ctagagattctggtgctgttaaaagagttgctgaattgaatggaaaaagatcttacttttgt
gatgaagatgaggaggagagagtgatggtggaggaagaaggtggttctgtttgtcaggtttt
gaattttgctgaagatgatgatgatgatgatgattatggtgaacaagggtaattgtggaaat
tggaattgatttgtttttgtggggttgtgtggaactggctatgttctgcttgattcttttgc
atttggtgtgaaactaaagatgaggtgaaaagtttatgcttgttaaattggattggtttat
atgttttgaaataataacaacaagcatgtgtcttgcttaataattgtatattgttttgtttg
ttttataatgatatggatttaatttgtatacacaatataatatagtatgcattgagagagtt
tttcgttcagtattcattctgattttagtgtttatctcattctagaagattgtattttgttg

SEQ ID NO 6: *Medicago truncalata* seedy1 protein
MNNTNNNNILLHSTQVQVWNNAAFDGEDFAMNSSSDSIKENLNPSAFNIVPSSNKRTIDDEI
AEIESEIKRLTSKLELLRVEKAERKIASEKRVSGIGTGRIVAAKFMEPKKNVTPKRNGVVFK
EETPKRNGVVSDTPKSRVNWRRGMSLGPMEIAGKVMAPPAMTITPATVNRRKSCFWKPQESC
EVMPSGITPATVNRRKSCFLKPQESCEENRRKTICKPNLNLNSNSVNSAVGSIKRVKKKDEE
IAQVQPKKLFEGEKSVKKSLKQGRIVASRYNSGGGGDARKRSFSENNKGLGSEIRAKKRWE
IPIEEVDVSGFVMLPKISTMRFVDESPRDSGAVKRVAELNGKRSYFCDEDEEERVMVEEEGG
SVCQVLNFAEDDDDDDDYGEQG

FIGURE 4 (continued)

SEQ ID NO 7: *Saccharum sp.* encoding sequence (partial 5' end)
Cgcaccgcgagtttcgaaaaccaacctatcgcgcctcagatcacgcgaggacgcgaggga
agcaggaatccctccgctcccagccgcctcctccgctcacccatcgatcgatcgtccgtccg
gtccaggggctctccggcggcggtggcgatggaggaggacccgctcatcccgctggtgcac
gtctggaacaacgccgccttcgaccacgcctcctcctccgcgtggcacgcccactcccctgt
gcccgcgagcgcacgtcgcgaggcggaggggacaaggagaaccaccgccccgaccccgacc
ccgacgtcgaggcggagatcggccacatcgaggcggagatcctgcgcctgtnctcccgcctn
caccaccttcgcacctccaagcagtcggagccgtccaagcgcggagaggtcgcgcccgcgcc
cgcggcgaaggcgaaagcggcggcggcggcgcggctgcggacgcggggctcagcctgggcc
cgctcgacgtcgccgctgccggtaaccccaacccgctcaccaccgacaaccagcagcagcag
ccgcgtgccgcgcagggtctgaagccgatcaagcaggccacggcggcggcgggcaagggcgt
aagacttgggccccttcgacatggtcggcgcgaaccctagggtccctccgcccn

SEQ ID NO 8: *Saccharum sp.* seedy1 protein (partial N-term)
MEEDPLIPLVHVWNNAAFDHASSSAWHAHSPVPASARREAEGDKENHRPDPDPDVEAEIGHI
EAEILRLXSRLHHLRTSKQSEPSKRGEVAPAPAAKAKAAAAARLRTRGLSLGPLDVAAAGNP
NPLTTDNQQQQPRAAQGLKPIKQATAAAGKGVRLGPLRHGRR

SEQ ID NO 9: *Zea Mays* seedy1 encoding sequence (partial 3'end)
ccacgcgtccggccgttcgagaggaggaaggccagcgttccaaggagcacgccgtccccgcc
agaccgtggccatccagcaatgccaggcacccactggatgccaggcaaggcaccgcagcaag
cagagccaaggcgaggagcgggagcataagccccagcaggttcaggaggcagtccacttcca
aggctgccgagacaagagcgggaaatgccaagcctacagaggcgacgaggggagggagcgaa
gcggtcaatcacaccagcaatgtagccacgacgaagaggccggcggggagctccaaggtcag
ggttgtcccgagccgctacagcatcccacctggctcctccctagcagctgtgacacaaggca
accgatgcaagcagtctctcccaggatcggctactgagaccagagtaaatctcactgagccg
ccgaacgacgagttgtctcctgaagaacttgccaaggttgcagagctgctcccaaggattag
gaccatgccgccttctgatgagagcccgcgtgactcgggatgtgccaagcgtgttgctgatt
tggtcgggaagcgatccttcttcactgctgcaggggacgatggcaatctcgttacgccctac
caggcacgggtggttgaacttgaatcacccgaggcagcagcagaagaagcagaagcttgaga
agtttgtctttgatcaattccgaagtggcttgcatctgggcgtggcctcttttgcagtgtg
tgctactacatagtctactgttacattcatatcatatcacatttcctatttttcccccttg
agacattgcttagtacttttgtgttgccttgtgaaaagagagtggaaggttcatctgctgat
nccttgtt

SEQ ID NO 10: *Zea Mays* seedy1 protein (partial C-term)
TRPAVREEEGQRSKEHAVPARPWPSSNARHPLDARQGTAASRAKARSGSISPSRFRRQSTSK
AAETRAGNAKPTEATRGGSEAVNHTSNVATTKRPAGSSKVRVVPSRYSIPPGSSLAAVTQGN
RCKQSLPGSATETRVNLTEPPNDELSPEELAKVAELLPRIRTMPPSDESPRDSGCAKRVADL
VGKRSFFTAAGDDGNLVTPYQARVVELESPEAAAEEAE

FIGURE 4 (continued)

SEQ ID NO 11: *Arabidospis thaliana* seedy1 coding sequence

atgacatcaattgaggcaacagaaacgcttaacgctcctccaaagcttcagatctggaacaa
cgctgccttcgacgatggagattctcaaatcacttccgccatcgaagcttcttcttggtctc
acctcaacgaatcattcgattccgattgtagcaaggagaatcagtttccgatttcggtttcc
tcttcgctccaatcctcagtctcgatcaccgaagctccgtcagcaaaatccaagaccgtgaa
gaccaaatccgccgcagatcggagtaaaaagcgagatatcgatgcagagatcgaagaagtag
agaaggagatcggacgattatcgacgaaattggagtcgctccgattagagaaggcggagcaa
accgcaagaagcattgctatacgtggaagaatcgttccggcgaagttcatggaatcatctca
gaaacaagtgaaattcgacgattcgtgttttacaggatcgaaatcaagagccactcgtagag
gcgttagtcttggaccagcggagatattcaattccgcgaagaaatctgaaactgtgactcct
cttcaatcagctcagaatcgacgcaagtcttgtttctttaagcttcctggaatcgaagaagg
tcaagtgacgacacgaggtaaaggaagaacgagtttgagtctgagtccgagatctcgcaaag
cgaaaatgacggcagctcagaagcaagcagctacgacggtggggtcaaagagagctgtgaag
aaagaagaaggagttctcttaacaatccagcctaagaggctattcaaagaagatgaaaagaa
tgtttctttaaggaaaccattgaaaccaggaagagttgtggctagtaggtacagtcaaatgg
gtaaaacgcagactggagagaaagatgttaggaaaaggtcgttgcctgaggatgaagagaaa
gagaatcataagaggtcggagaagagaagagcttctgatgaaagtaacaagagtgaagggag
agtgaagaagagatgggagattccaagtgaagttgatctgtatagcagtggtgagaacggtg
acgagtctcctatagttaaggagctacctaagatcagaacgcttcgtcgtgtgggagggagc
cctcgtgattcaggtgctgctaagagagttgcagaattacaagccaaggatcgtaacttcac
tttttgccagcttctgaagtttgaagaatgaatgatccgcttatcaatttgagtaaaatcca
caactcttgttgtggtt

SEQ ID NO 12: *Arabidospis thaliana* seedy1 protein

MTSIEATETLNAPPKLQIWNNAAFDDGDSQITSAIEASSWSHLNESFDSDCSKENQFPISVS
SSLQSSVSITEAPSAKSKTVKTKSAADRSKKRDIDAEIEEVEKEIGRLSTKLESLRLEKAEQ
TARSIAIRGRIVPAKFMESSQKQVKFDDSCFTGSKSRATRRGVSLGPAEIFNSAKKSETVTP
LQSAQNRRKSCFFKLPGIEEGQVTTRGKGRTSLSLSPRSRKAKMTAAQKQAATTVGSKRAVK
KEEGVLLTIQPKRLFKEDEKNVSLRKPLKPGRVVASRYSQMGKTQTGEKDVRKRSLPEDEEK
ENHKRSEKRRASDESNKSEGRVKKRWEIPSEVDLYSSGENGDESPIVKELPKIRTLRRVGGS
PRDSGAAKRVAELQAKDRNFTFCQLLKFEE

FIGURE 4 (continued)

SEQ ID NO 13: Sequence of the [PRO0090 - CDS0689 - terminator] expression cassette

cttctacatcggcttaggtgtagcaacacgactttattattattattattattattatt
attttacaaaaatataaaatagatcagtccctcaccacaagtagagcaagttggtgagttat
tgtaaagttctacaaagctaatttaaaagttattgcattaacttatttcatattacaaacaa
gagtgtcaatggaacaatgaaaaccatatgacatactataattttgtttttattattgaaat
tatataattcaaagagaataaatccacatagccgtaaagttctacatgtggtgcattaccaa
aatatatatagcttacaaaacatgacaagcttagtttgaaaaattgcaatccttatcacatt
gacacataaagtgagtgatgagtcataatattattttctttgctacccatcatgtatatatg
atagccacaaagttactttgatgatgatatcaaagaacattttaggtgcacctaacagaat
atccaaataatatgactcacttagatcataatagagcatcaagtaaaactaacactctaaag
caaccgatgggaaagcatctataaatagacaagcacaatgaaaatcctcatcatccttcacc
acaattcaaatattatagttgaagcatagtagtaatttaaatcaactagggatatcacaagt
ttgtacaaaaaagcaggctggtaccggtccggaattcccgggatatcgtcgacccacgcgtc
cgctgacgcgtgggttccactacatcaagacatctactacactcatctttttgcacttatt
gggtgtaaatttttgaaacccagttgagaaaaatgagtgtgttacaatacccagaagggatt
gacccagcagatgttcagatatggaacaatgcagcatttgataatggagattctgaagattt
gtcttcgctgaaacgttcttggtctcctctgaaacccctttcggttaggccatcagattcct
ttgaatctgatttgtcaagtaaggaaaatcaaactcctttatttgagaattcatctgttaat
ctctcatctccgttacccataaagccacttaaccctaatggggctctggaaaattcaagact
caagccgaacaagcccaattccaaacagagtcttgatgagatggcggctagaaagagcggaa
agggaaatgatttccgtgatgagaagaaaatagacgaggaaattgaagaaattcagatggag
attagtaggttgagttcaagattagaggctttgagaattgaaaaggctgagaaaactgttgc
taagactgttgaaaagcgaggaagggttgtggcagcaaagtttatggagccaaaacaaagtg
ttattaagattgaagagcgtatatcaatgagtgcaagaacaaaggtggagcagagaaggggt
cttagtttaggaccatctgagattttactggaacgcggcggcgagggttgagtatggggcc
atcagatattctagcagggacaacaaaggcacggcaattgggaaagcaagagatgattatta
ctcctattcagccaatacaaaacaggcgaaagtcgtgtttttggaagcttcaagagattgaa
gaagagggaaaaagttcaagccttagtcctaaatcaagaaaaactgctgcaagaacaatggt
tacaacaaggcaggcagttactacaattgcatcaaagaagaatttgaaaaaagatgatggac
ttttgagttcagttcagccaaagaagttgtttaaagatctcgaaaagtctgctgctgctaat
aagaagccccagaggccggggagggttgtggctagtaggtataatcagagtacaattcagtc
atcagtagtgagaaagaggtctttacctgaaaatgataaggatgagagtaagagaaatgata
agaaacggtcgttatctgtagggaaaacgcgtgtgtctcaaactgagagcaagaatttgggt
actgaaagtagggtgaaaaagagatgggaaattcctagtgagattgtagttcatggaaacac
agagagtgagaaatctccactaagcattattgtgaagcctgatttgcttccgcgaattagga
ttgctcggtgtgtgaatgagactcttagggattctggacctgctaaaagaatgatagagttg
ataggcaagaaatcgtttttcagtagtgatgaagataaggagccacctgtctgtcaagtttt
aagttttgcagaggaagatgctgaagaggaataatgtgtaataaagggagctgctaactctt
ttcatgctctttcaattttcaatcctgccttttaattttgttcattcgtgccttttaattg
aatggggaagcattcttttgcttcctcaaactggtattctagcttctgaattacattgtatg
gtacaatatgaataaggttttgtcttccggcaggttgtccaagttagtttttagcttaaaat
agatgcggcagcggccgctctagagtatccctcgagggcccaagcttacgcgtacccagct

FIGURE 4 (continued)

ttcttgtacaaagtggtgatatcacaagcccgggcggtcttctagggataacagggtaatta
tatccctctagatcacaagcccgggcggtcttctacgatgattgagtaataatgtgtcacgc
atcaccatgggtggcagtgtcagtgtgagcaatgacctgaatgaacaattgaaatgaaaaga
aaaaaagtactccatctgttccaaattaaaattcatttaaccttttaataggtttatacaa
taattgatatatgttttctgtatatgtctaatttgttatcatccgggcggtcttctagggat
aacagggtaattatatccctctagacaacacacaacaaataagagaaaaaacaaataatatt
aatttgagaatgaacaaaaggaccatatcattcattaactcttctccatccatttccatttc
acagttcgatagcgaaaaccgaataaaaaacacagtaaattacaagcacaacaaatggtaca
agaaaaacagttttcccaatgccataatactcgaac

SEQ ID NO 14: rice prolamin RP6 promoter sequence
ccttctacatcggcttaggtgtagcaacacgactttattattattattattattattattat
tattttacaaaaatataaaatagatcagtccctcaccacaagtagagcaagttggtgagtta
ttgtaaagttctacaaagctaatttaaaagttattgcattaacttatttcatattacaaaca
agagtgtcaatggaacaatgaaaaccatatgacatactataatttttgtttttattattgaaa
ttatataattcaaagagaataaatccacatagccgtaaagttctacatgtggtgcattacca
aaatatatatagcttacaaaacatgacaagcttagtttgaaaaattgcaatccttatcacat
tgacacataaagtgagtgatgagtcataatattatttttcttgctacccatcatgtatatat
gatagccacaaagttactttgatgatgatatcaaagaacattttaggtgcacctaacagaa
tatccaaataatatgactcacttagatcataatagagcatcaagtaaaactaacactctaaa
gcaaccgatgggaaagcatctataaatagacaagcacaatgaaaatcctcatcatccttcac
cacaattcaaatattatagttgaagcatagtagtagaatccaacaaca

SEQ ID NO 15: Motif 1 CORE SEQUENCE
WXNAXXD

SEQ ID NO 16: Motif 2 CORE SEQUENCE
KENXXP

SEQ ID NO 17: Motif 3 (coiled coil) CORE SEQUENCE
$EX_{1-6}EXXRLXXXLXXLR$

SEQ ID NO 18: Motif 4 CORE SEQUENCE
$LPXIX_{1-10}RDSGXXKRX_{1-6}K$

FIGURE 4 (continued)

SEQUENCE LISTING

<110> CropDesign N.V.

<120> Seedy1 sequence for making plants having changed growth characteristics

<130> 105-Seedy1-EP

<160> 18

<170> PatentIn version 3.1

<210> 1

<211> 1428

<212> DNA

<213> Nicotiana tabacum

<220>

<221> misc_feature

<223> seedy1 coding sequence (CDS0689)

<400> 1
atgagtgtgt tacaataccc agaagggatt gacccagcag atgttcagat atggaacaat 60

gcagcatttg ataatggaga ttctgaagat ttgtcttcgc tgaaacgttc ttggtctcct 120

ctgaaacccc tttcggttag gccatcagat tcctttgaat ctgatttgtc aagtaaggaa 180

aatcaaactc ctttatttga gaattcatct gttaatctct catctccgtt acccataaag 240

ccacttaacc ctaatggggc tctggaaaat tcaagactca agccgaacaa gcccaattcc 300

aaacagagtc ttgatgagat ggcggctaga aagagcggaa agggaaatga tttccgtgat 360

gagaagaaaa tagacgagga aattgaagaa attcagatgg agattagtag gttgagttca 420

agattagagg ctttgagaat tgaaaaggct gagaaaactg ttgctaagac tgttgaaaag 480

```
cgaggaaggg ttgtggcagc aaagtttatg gagccaaaac aaagtgttat taagattgaa     540
gagcgtatat caatgagtgc aagaacaaag gtggagcaga gaaggggtct tagtttagga     600
ccatctgaga tttttactgg aacgcggcgg cgagggttga gtatggggcc atcagatatt     660
ctagcaggga caacaaaggc acggcaattg ggaaagcaag agatgattat tactcctatt     720
cagccaatac aaaacaggcg aaagtcgtgt ttttggaagc ttcaagagat tgaagaagag     780
ggaaaaagtt caagccttag tcctaaatca agaaaaactg ctgcaagaac aatggttaca     840
acaaggcagg cagttactac aattgcatca aagaagaatt tgaaaaaaga tgatggactt     900
ttgagttcag ttcagccaaa gaagttgttt aaagatctcg aaaagtctgc tgctgctaat     960
aagaagcccc agaggccggg gagggttgtg gctagtaggt ataatcagag tacaattcag    1020
tcatcagtag tgagaaagag gtctttacct gaaaatgata aggatgagag taagagaaat    1080
gataagaaac ggtcgttatc tgtagggaaa acgcgtgtgt ctcaaactga gagcaagaat    1140
ttgggtactg aaagtagggt gaaaaagaga tgggaaattc ctagtgagat tgtagttcat    1200
ggaaacacag agagtgagaa atctccacta agcattattg tgaagcctga tttgcttccg    1260
cgaattagga ttgctcggtg tgtgaatgag actcttaggg attctggacc tgctaaaaga    1320
atgatagagt tgataggcaa gaaatcgttt ttcagtagtg atgaagataa ggagccacct    1380
gtctgtcaag ttttaagttt tgcagaggaa gatgctgaag aggaataa                 1428
```

<210> 2

<211> 475

<212> PRT

<213> Nicotiana tabacum

<220>

<221> MISC_FEATURE

<223> seedy1 protein (CDS0689)

<400> 2

```
Met Ser Val Leu Gln Tyr Pro Glu Gly Ile Asp Pro Ala Asp Val Gln
1               5                   10                  15

Ile Trp Asn Asn Ala Ala Phe Asp Asn Gly Asp Ser Glu Asp Leu Ser
            20                  25                  30
```

```
Ser Leu Lys Arg Ser Trp Ser Pro Leu Lys Pro Leu Ser Val Arg Pro
        35                  40                  45

Ser Asp Ser Phe Glu Ser Asp Leu Ser Ser Lys Glu Asn Gln Thr Pro
    50                  55                  60

Leu Phe Glu Asn Ser Ser Val Asn Leu Ser Ser Pro Leu Pro Ile Lys
65                  70                  75                  80

Pro Leu Asn Pro Asn Gly Ala Leu Glu Asn Ser Arg Leu Lys Pro Asn
                85                  90                  95

Lys Pro Asn Ser Lys Gln Ser Leu Asp Glu Met Ala Ala Arg Lys Ser
            100                 105                 110

Gly Lys Gly Asn Asp Phe Arg Asp Glu Lys Lys Ile Asp Glu Glu Ile
        115                 120                 125

Glu Glu Ile Gln Met Glu Ile Ser Arg Leu Ser Ser Arg Leu Glu Ala
    130                 135                 140

Leu Arg Ile Glu Lys Ala Glu Lys Thr Val Ala Lys Thr Val Glu Lys
145                 150                 155                 160

Arg Gly Arg Val Val Ala Ala Lys Phe Met Glu Pro Lys Gln Ser Val
                165                 170                 175

Ile Lys Ile Glu Glu Arg Ile Ser Met Ser Ala Arg Thr Lys Val Glu
            180                 185                 190

Gln Arg Arg Gly Leu Ser Leu Gly Pro Ser Glu Ile Phe Thr Gly Thr
        195                 200                 205

Arg Arg Arg Gly Leu Ser Met Gly Pro Ser Asp Ile Leu Ala Gly Thr
    210                 215                 220

Thr Lys Ala Arg Gln Leu Gly Lys Gln Glu Met Ile Ile Thr Pro Ile
225                 230                 235                 240

Gln Pro Ile Gln Asn Arg Arg Lys Ser Cys Phe Trp Lys Leu Gln Glu
                245                 250                 255

Ile Glu Glu Glu Gly Lys Ser Ser Ser Leu Ser Pro Lys Ser Arg Lys
```

260 265 270

Thr Ala Ala Arg Thr Met Val Thr Thr Arg Gln Ala Val Thr Thr Ile
275 280 285

Ala Ser Lys Lys Asn Leu Lys Lys Asp Asp Gly Leu Leu Ser Ser Val
290 295 300

Gln Pro Lys Lys Leu Phe Lys Asp Leu Glu Lys Ser Ala Ala Ala Asn
305 310 315 320

Lys Lys Pro Gln Arg Pro Gly Arg Val Val Ala Ser Arg Tyr Asn Gln
325 330 335

Ser Thr Ile Gln Ser Ser Val Val Arg Lys Arg Ser Leu Pro Glu Asn
340 345 350

Asp Lys Asp Glu Ser Lys Arg Asn Asp Lys Lys Arg Ser Leu Ser Val
355 360 365

Gly Lys Thr Arg Val Ser Gln Thr Glu Ser Lys Asn Leu Gly Thr Glu
370 375 380

Ser Arg Val Lys Lys Arg Trp Glu Ile Pro Ser Glu Ile Val Val His
385 390 395 400

Gly Asn Thr Glu Ser Glu Lys Ser Pro Leu Ser Ile Ile Val Lys Pro
405 410 415

Asp Leu Leu Pro Arg Ile Arg Ile Ala Arg Cys Val Asn Glu Thr Leu
420 425 430

Arg Asp Ser Gly Pro Ala Lys Arg Met Ile Glu Leu Ile Gly Lys Lys
435 440 445

Ser Phe Phe Ser Ser Asp Glu Asp Lys Glu Pro Pro Val Cys Gln Val
450 455 460

Leu Ser Phe Ala Glu Glu Asp Ala Glu Glu Glu
465 470 475

<210> 3

<211> 1336

<212> DNA

<213> Oryza sativa

<220>

<221> misc_feature

<223> seedy1 coding sequence

<400> 3
atggaggagg acccgctcat cccgctggtc cacgtctgga acaacgccgc cttcgacgac 60

tcctcgtgtt ccagatcggc ttggctcccc caaagccccg ccgtcgcggc cgtccgcaag 120

ggcgacaagg agaatcaccg ccccgaggtt gttgatgtcg ccgccggcta cgacgtcgag 180

gccgagatcg gccacatcga ggcggagatc ctgcgcctct cgtcccggct ccaccatctc 240

cgcgtctcca agcagccgga gcccaaccgc gacgacgctc cgatggggga gatggtcgcg 300

aaggtgaggc cccggccgag gggcctcagc ctcgggcccc tggatgtgat ctccatcgtc 360

aatcgtgaga agcatccgct gcgcaccaag cagcctccgg cgacgcgggg caggggggctc 420

agcctcgggc ccatggagat cgccgcggcg aaccctaggg tgcccgcggc ggcgcagcat 480

cagcaacagc aacgcgctgg cacggcgcgg atcctgaagc caatcaagga gcctccggtg 540

cagcgtcgca ggggcgtcag cctcgggccg ttggagatcc accacggcgt cggcagcaag 600

gcaccagcgg cggcgcgagc caagccgttc accaccaagc tcaacgccat tcgagaagaa 660

acccgacccct ccaagcaatt cgccgtcccc gccaagccat ggccgtcgag caatacaagg 720

cagacactgg actcgaggca aggaacagca gcaagtcgag cgaaggcgag gagcccgagc 780

cccaggccca ggaggcaatc caatggcaag gctactgaca caaggggagg caacaaggtg 840

gtggatgagc tcaagcccaa aggtgcgtcg tcaagtcaga gcggcagcgc cgccgccgcc 900

gccactgcca agaggatggc ggggagctcc aagatgaggg tcatcccgag ccgctacagc 960

ctcactcctg gcgcttccct tggaagcagt ggagcacagg agaggcgacg caagcagtct 1020

ctcccaggat catcagggga tgcgaaccag aatgaggaaa tcagagcgaa ggtcatcgag 1080

ccttccaatg atccactctc tcctcaaacg atctccaagg ttgctgaaat gctcccaaag 1140

atcaggacca tgccgcctcc tgacgagagc cctcgcgatt ccggatgcgc caagcgggtt 1200

gccgaattgg tcgggaagcg ctcgttcttc acggctgcag ccgaggacgg gcgggcgctc 1260

gacgtcgaag cacccgaggc ggtcgcagaa gcttgagatg aaccaccatg gtttgatccg 1320

ttccttccat cagctc 1336

<210> 4

<211> 431

<212> PRT

<213> Oryza sativa

<220>

<221> MISC_FEATURE

<223> seedy1 protein

<400> 4

Met Glu Glu Asp Pro Leu Ile Pro Leu Val His Val Trp Asn Asn Ala
1 5 10 15

Ala Phe Asp Asp Ser Ser Cys Ser Arg Ser Ala Trp Leu Pro Gln Ser
20 25 30

Pro Ala Val Ala Ala Val Arg Lys Gly Asp Lys Glu Asn His Arg Pro
35 40 45

Glu Val Val Asp Val Ala Ala Gly Tyr Asp Val Glu Ala Glu Ile Gly
50 55 60

His Ile Glu Ala Glu Ile Leu Arg Leu Ser Ser Arg Leu His His Leu
65 70 75 80

Arg Val Ser Lys Gln Pro Glu Pro Asn Arg Asp Asp Ala Pro Met Gly
85 90 95

Glu Met Val Ala Lys Val Arg Pro Arg Pro Arg Gly Leu Ser Leu Gly
100 105 110

Pro Leu Asp Val Ile Ser Ile Val Asn Arg Glu Lys His Pro Leu Arg
115 120 125

Thr Lys Gln Pro Pro Ala Thr Arg Gly Arg Gly Leu Ser Leu Gly Pro
130 135 140

```
Met Glu Ile Ala Ala Ala Asn Pro Arg Val Pro Ala Ala Ala Gln His
145                 150                 155                 160

Gln Gln Gln Gln Arg Ala Gly Thr Ala Arg Ile Leu Lys Pro Ile Lys
                165                 170                 175

Glu Pro Pro Val Gln Arg Arg Arg Gly Val Ser Leu Gly Pro Leu Glu
            180                 185                 190

Ile His His Gly Val Gly Ser Lys Ala Pro Ala Ala Ala Arg Ala Lys
        195                 200                 205

Pro Phe Thr Thr Lys Leu Asn Ala Ile Arg Glu Glu Thr Arg Pro Ser
    210                 215                 220

Lys Gln Phe Ala Val Pro Ala Lys Pro Trp Pro Ser Ser Asn Thr Arg
225                 230                 235                 240

Gln Thr Leu Asp Ser Arg Gln Gly Thr Ala Ala Ser Arg Ala Lys Ala
                245                 250                 255

Arg Ser Pro Ser Pro Arg Pro Arg Arg Gln Ser Asn Gly Lys Ala Thr
            260                 265                 270

Asp Thr Arg Gly Gly Asn Lys Val Val Asp Glu Leu Lys Pro Lys Gly
        275                 280                 285

Ala Ser Ser Ser Gln Ser Gly Ser Ala Ala Ala Ala Ala Thr Ala Lys
    290                 295                 300

Arg Met Ala Gly Ser Ser Lys Met Arg Val Ile Pro Ser Arg Tyr Ser
305                 310                 315                 320

Leu Thr Pro Gly Ala Ser Leu Gly Ser Ser Gly Ala Gln Glu Arg Arg
                325                 330                 335

Arg Lys Gln Ser Leu Pro Gly Ser Ser Gly Asp Ala Asn Gln Asn Glu
            340                 345                 350

Glu Ile Arg Ala Lys Val Ile Glu Pro Ser Asn Asp Pro Leu Ser Pro
        355                 360                 365

Gln Thr Ile Ser Lys Val Ala Glu Met Leu Pro Lys Ile Arg Thr Met
    370                 375                 380
```

Pro Pro Pro Asp Glu Ser Pro Arg Asp Ser Gly Cys Ala Lys Arg Val
385 390 395 400

Ala Glu Leu Val Gly Lys Arg Ser Phe Phe Thr Ala Ala Ala Glu Asp
405 410 415

Gly Arg Ala Leu Asp Val Glu Ala Pro Glu Ala Val Ala Glu Ala
420 425 430

<210> 5

<211> 1860

<212> DNA

<213> Medicago truncculata

<220>

<221> misc_feature

<223> seedy1 coding sequence

<400> 5
aaaaacgtta aggactaaaa atataataaa atttaagtag ggattcataa tggaagcacc 60

cctatttaca gggatcttaa atataattaa ccctaatatt tatgacagaa accctttttga 120

aatcacatcg gagcgtgtat gagtagccgt ttcacatcca acggccagta agagcgtaac 180

tttatttctt ccctcttcaa tctccaacgg tcacataatc tcttccaaat acaaataatt 240

ccctctttca acctcactct tcatttcttc aacccaaacc caaaaaacta atcagattct 300

tcttaaatct tgaaaccttt ctcccaaaag cacttaaata aaaaagcact taaccatgaa 360

taacacaaac aacaacaaca ttcttcttca ttccacacag gttcaagtgt ggaacaacgc 420

agcattcgat ggtgaagatt tcgccatgaa ttcatcttct gattccatca aagagaatct 480

aaacccatcc gcattcaaca ttgttccttc ttcaaacaaa agaactattg atgatgaaat 540

tgcggaaatt gaaagtgaaa ttaagcgatt aacttcgaag ctggaattgc ttcgtgttga 600

aaaagctgaa agaaaaatcg cttctgaaaa gcgtgttagt ggaattggta ctggaagaat 660

agtagcagcg aagtttatgg aaccgaagaa aaacgttaca ccgaaacgaa acggtgtcgt 720

tttcaaggag gagacaccga aacgaaacgg tgtcgtttcg gatacgccga aatctagggt 780

taattggaga agagggatga gtttaggtcc gatggagatt gccgggaaag tgatggcacc 840

gccggcgatg acgattactc cggcgacggt gaatcggagg aagtcttgtt tctggaaacc 900

gcaggaaagt tgtgaagtaa tgccgtcggg gattactccg gcgacggtga ataggaggaa 960

atcttgtttt ttgaaacctc aagaaagttg tgaagaaaat cgaagaaaaa cgatttgcaa 1020

accgaatttg aatttgaatt caaattcagt taattctgcg gttggatcga ttaagcgtgt 1080

gaagaagaaa gatgaagaaa ttgctcaggt tcaaccgaag aagctgtttg aaggtgaaaa 1140

atcagtgaag aaatcgttga aacaaggtag aattgttgca agccggtata attccggtgg 1200

tggtggtggt gatgcgagga aaagatcgtt ttcggagaat aataagggtt tagggagtga 1260

aatcagggct aagaagagat gggagatacc aattgaagaa gtggatgtga gtggttttgt 1320

tatgttaccg aagatttcga caatgaggtt tgttgatgag agtcctagag attctggtgc 1380

tgttaaaaga gttgctgaat tgaatggaaa aagatcttac ttttgtgatg aagatgagga 1440

ggagagagtg atggtggagg aagaaggtgg ttctgtttgt caggttttga attttgctga 1500

agatgatgat gatgatgatg attatggtga acaagggtaa ttgtggaaat tggaattgat 1560

ttgtttttgt ggggttgtgt ggaactggct atgttctgct tgattctttt gcattttggt 1620

gtgaaactaa agatgaggtg aaaagtttat gcttgttaaa ttggattggt ttatatgttt 1680

tgaaataata acaacaagca tgtgtcttgc ttaataattg tatattgttt tgtttgtttt 1740

ataatgatat ggatttaatt tgtatacaca atataatata gtatgcattg agagagtttt 1800

tcgttcagta ttcattctga ttttagtgtt tatctcattc tagaagattg tattttgttg 1860

<210> 6

<211> 394

<212> PRT

<213> Medicago trunculata

<220>

<221> MISC_FEATURE

<223> seedy1 protein

<400> 6

Met Asn Asn Thr Asn Asn Asn Asn Ile Leu Leu His Ser Thr Gln Val
1 5 10 15

Gln Val Trp Asn Asn Ala Ala Phe Asp Gly Glu Asp Phe Ala Met Asn
20 25 30

Ser Ser Ser Asp Ser Ile Lys Glu Asn Leu Asn Pro Ser Ala Phe Asn
35 40 45

Ile Val Pro Ser Ser Asn Lys Arg Thr Ile Asp Asp Glu Ile Ala Glu
50 55 60

Ile Glu Ser Glu Ile Lys Arg Leu Thr Ser Lys Leu Glu Leu Leu Arg
65 70 75 80

Val Glu Lys Ala Glu Arg Lys Ile Ala Ser Glu Lys Arg Val Ser Gly
85 90 95

Ile Gly Thr Gly Arg Ile Val Ala Ala Lys Phe Met Glu Pro Lys Lys
100 105 110

Asn Val Thr Pro Lys Arg Asn Gly Val Val Phe Lys Glu Glu Thr Pro
115 120 125

Lys Arg Asn Gly Val Val Ser Asp Thr Pro Lys Ser Arg Val Asn Trp
130 135 140

Arg Arg Gly Met Ser Leu Gly Pro Met Glu Ile Ala Gly Lys Val Met
145 150 155 160

Ala Pro Pro Ala Met Thr Ile Thr Pro Ala Thr Val Asn Arg Arg Lys
165 170 175

Ser Cys Phe Trp Lys Pro Gln Glu Ser Cys Glu Val Met Pro Ser Gly
180 185 190

Ile Thr Pro Ala Thr Val Asn Arg Arg Lys Ser Cys Phe Leu Lys Pro
195 200 205

Gln Glu Ser Cys Glu Glu Asn Arg Arg Lys Thr Ile Cys Lys Pro Asn
210 215 220

Leu Asn Leu Asn Ser Asn Ser Val Asn Ser Ala Val Gly Ser Ile Lys
225 230 235 240

Arg Val Lys Lys Lys Asp Glu Glu Ile Ala Gln Val Gln Pro Lys Lys
245 250 255

Leu Phe Glu Gly Glu Lys Ser Val Lys Lys Ser Leu Lys Gln Gly Arg
260 265 270

Ile Val Ala Ser Arg Tyr Asn Ser Gly Gly Gly Gly Gly Asp Ala Arg
275 280 285

Lys Arg Ser Phe Ser Glu Asn Asn Lys Gly Leu Gly Ser Glu Ile Arg
290 295 300

Ala Lys Lys Arg Trp Glu Ile Pro Ile Glu Glu Val Asp Val Ser Gly
305 310 315 320

Phe Val Met Leu Pro Lys Ile Ser Thr Met Arg Phe Val Asp Glu Ser
325 330 335

Pro Arg Asp Ser Gly Ala Val Lys Arg Val Ala Glu Leu Asn Gly Lys
340 345 350

Arg Ser Tyr Phe Cys Asp Glu Asp Glu Glu Glu Arg Val Met Val Glu
355 360 365

Glu Glu Gly Gly Ser Val Cys Gln Val Leu Asn Phe Ala Glu Asp Asp
370 375 380

Asp Asp Asp Asp Asp Tyr Gly Glu Gln Gly
385 390

<210> 7

<211> 674

<212> DNA

<213> Saccharum sp.

<220>

<221> misc_feature

<223> seedy1 coding sequence (partial 5' end)

<220>

<221> misc_feature

<222> (362)..(362)

<223> n can be a, c, g or t

<220>

<221> misc_feature

<222> (372)..(372)

<223> n can be a, c, g or t

<220>

<221> misc_feature

<222> (674)..(674)

<223> n can be a, c, g or t

<400> 7
cgcaccgcga gtttcgaaaa accaacctat cgcgcctcag atcacgcgag gacgcgaggg 60

gaagcaggaa tccctccgct cccagccgcc tcctccgctc acccatcgat cgatcgtccg 120

tccggtccag gggctctcc ggcggcggtg gcgatggagg aggacccgct catcccgctg 180

gtgcacgtct ggaacaacgc cgccttcgac cacgcctcct cctccgcgtg gcacgcccac 240

tcccctgtgc ccgcgagcgc acgtcgcgag gcggaggggg acaaggagaa ccaccgcccc 300

gaccccgacc ccgacgtcga ggcggagatc ggccacatcg aggcggagat cctgcgcctg 360

tnctcccgcc tncaccacct tcgcacctcc aagcagtcgg agccgtccaa gcgcggagag 420

gtcgcgcccg cgcccgcggc gaaggcgaaa gcggcggcgg cggcgcggct gcggacgcgg 480

gggctcagcc tgggcccgct cgacgtcgcc gctgccggta accccaaccc gctcaccacc 540

gacaaccagc agcagcagcc gcgtgccgcg cagggtctga agccgatcaa gcaggccacg 600

gcggcggcgg gcaagggcgt aagacttggg ccccttcgac atggtcggcg cgaaccctag 660

ggtccctccg cccn 674

<210> 8

<211> 166

<212> PRT

<213> Saccharum sp.

<220>

<221> MISC_FEATURE

<223> seedy1 protein

<220>

<221> MISC_FEATURE

<223> seedy1 protein (partial N term)

<220>

<221> MISC_FEATURE

<222> (70)..(70)

<223> Xaa can be any amino acid

<400> 8

```
Met Glu Glu Asp Pro Leu Ile Pro Leu Val His Val Trp Asn Asn Ala
1               5                   10                  15

Ala Phe Asp His Ala Ser Ser Ser Ala Trp His Ala His Ser Pro Val
            20                  25                  30

Pro Ala Ser Ala Arg Arg Glu Ala Glu Gly Asp Lys Glu Asn His Arg
        35                  40                  45

Pro Asp Pro Asp Pro Asp Val Glu Ala Glu Ile Gly His Ile Glu Ala
    50                  55                  60

Glu Ile Leu Arg Leu Xaa Ser Arg Leu His His Leu Arg Thr Ser Lys
65                  70                  75                  80

Gln Ser Glu Pro Ser Lys Arg Gly Glu Val Ala Pro Ala Pro Ala Ala
                85                  90                  95

Lys Ala Lys Ala Ala Ala Ala Ala Arg Leu Arg Thr Arg Gly Leu Ser
            100                 105                 110
```

Leu Gly Pro Leu Asp Val Ala Ala Ala Gly Asn Pro Asn Pro Leu Thr
        115   120   125

Thr Asp Asn Gln Gln Gln Gln Pro Arg Ala Ala Gln Gly Leu Lys Pro
    130   135   140

Ile Lys Gln Ala Thr Ala Ala Ala Gly Lys Gly Val Arg Leu Gly Pro
145   150   155   160

Leu Arg His Gly Arg Arg
                165

<210> 9

<211> 876

<212> DNA

<213> Zea Mays

<220>

<221> misc_feature

<223> seedy1 coding sequence (partila 3' end)

<220>

<221> misc_feature

<222> (869)..(869)

<223> n = a, c, g or t

<400> 9
ccacgcgtcc ggccgttcga gaggaggaag gccagcgttc caaggagcac gccgtccccg 60

ccagaccgtg gccatccagc aatgccaggc acccactgga tgccaggcaa ggcaccgcag 120

caagcagagc caaggcgagg agcgggagca taagccccag caggttcagg aggcagtcca 180

cttccaaggc tgccgagaca agagcgggaa atgccaagcc tacagaggcg acgaggggag 240

ggagcgaagc ggtcaatcac accagcaatg tagccacgac gaagaggccg gcggggagct 300

ccaaggtcag ggttgtcccg agccgctaca gcatcccacc tggctcctcc ctagcagctg 360

tgacacaagg caaccgatgc aagcagtctc tcccaggatc ggctactgag accagagtaa 420

atctcactga gccgccgaac gacgagttgt ctcctgaaga acttgccaag gttgcagagc 480

tgctcccaag gattaggacc atgccgcctt ctgatgagag cccgcgtgac tcgggatgtg 540

ccaagcgtgt tgctgatttg gtcgggaagc gatccttctt cactgctgca ggggacgatg 600

gcaatctcgt tacgccctac caggcacggg tggttgaact tgaatcaccc gaggcagcag 660

cagaagaagc agaagcttga gaagtttgtc tttgatcaat tccgaagtgg cttgcatctg 720

ggcgtggcct ctttttgcag tgtgtgctac tacatagtct actgttacat tcatatcata 780

tcacatttcc tattttttcc cccttgagac attgcttagt acttttgtgt tgccttgtga 840

aaagagagtg gaaggttcat ctgctgatnc cttgtt 876

<210> 10

<211> 224

<212> PRT

<213> Zea Mays

<220>

<221> MISC_FEATURE

<223> seedy1 protein (partial C term)

<400> 10

Thr Arg Pro Ala Val Arg Glu Glu Glu Gly Gln Arg Ser Lys Glu His
1   5   10   15

Ala Val Pro Ala Arg Pro Trp Pro Ser Ser Asn Ala Arg His Pro Leu
            20   25   30

Asp Ala Arg Gln Gly Thr Ala Ala Ser Arg Ala Lys Ala Arg Ser Gly
        35   40   45

Ser Ile Ser Pro Ser Arg Phe Arg Arg Gln Ser Thr Ser Lys Ala Ala
    50   55   60

Glu Thr Arg Ala Gly Asn Ala Lys Pro Thr Glu Ala Thr Arg Gly Gly
65   70   75   80

Ser Glu Ala Val Asn His Thr Ser Asn Val Ala Thr Thr Lys Arg Pro

85 90 95

Ala Gly Ser Ser Lys Val Arg Val Val Pro Ser Arg Tyr Ser Ile Pro
100 105 110

Pro Gly Ser Ser Leu Ala Ala Val Thr Gln Gly Asn Arg Cys Lys Gln
115 120 125

Ser Leu Pro Gly Ser Ala Thr Glu Thr Arg Val Asn Leu Thr Glu Pro
130 135 140

Pro Asn Asp Glu Leu Ser Pro Glu Glu Leu Ala Lys Val Ala Glu Leu
145 150 155 160

Leu Pro Arg Ile Arg Thr Met Pro Pro Ser Asp Glu Ser Pro Arg Asp
165 170 175

Ser Gly Cys Ala Lys Arg Val Ala Asp Leu Val Gly Lys Arg Ser Phe
180 185 190

Phe Thr Ala Ala Gly Asp Asp Gly Asn Leu Val Thr Pro Tyr Gln Ala
195 200 205

Arg Val Val Glu Leu Glu Ser Pro Glu Ala Ala Ala Glu Glu Ala Glu
210 215 220

<210> 11

<211> 1257

<212> DNA

<213> Arabidospis thaliana

<220>

<221> misc_feature

<223> seedy1 coding sequence

<400> 11
atgacatcaa ttgaggcaac agaaacgctt aacgctcctc caaagcttca gatctggaac 60

aacgctgcct tcgacgatgg agattctcaa atcacttccg ccatcgaagc ttcttcttgg 120

tctcacctca acgaatcatt cgattccgat tgtagcaagg agaatcagtt tccgatttcg 180

gtttcctctt cgctccaatc ctcagtctcg atcaccgaag ctccgtcagc aaaatccaag 240

accgtgaaga ccaaatccgc cgcagatcgg agtaaaaagc gagatatcga tgcagagatc 300

gaagaagtag agaaggagat cggacgatta tcgacgaaat tggagtcgct ccgattagag 360

aaggcggagc aaaccgcaag aagcattgct atacgtggaa gaatcgttcc ggcgaagttc 420

atggaatcat ctcagaaaca agtgaaattc gacgattcgt gttttacagg atcgaaatca 480

agagccactc gtagaggcgt tagtcttgga ccagcggaga tattcaattc cgcgaagaaa 540

tctgaaactg tgactcctct tcaatcagct cagaatcgac gcaagtcttg tttctttaag 600

cttcctggaa tcgaagaagg tcaagtgacg acacgaggta aaggaagaac gagtttgagt 660

ctgagtccga gatctcgcaa agcgaaaatg acggcagctc agaagcaagc agctacgacg 720

gtggggtcaa agagagctgt gaagaaagaa gaaggagttc tcttaacaat ccagcctaag 780

aggctattca aagaagatga aaagaatgtt tctttaagga aaccattgaa accaggaaga 840

gttgtggcta gtaggtacag tcaaatgggt aaaacgcaga ctggagagaa agatgttagg 900

aaaaggtcgt tgcctgagga tgaagagaaa gagaatcata agaggtcgga gaagagaaga 960

gcttctgatg aaagtaacaa gagtgaaggg agagtgaaga agagatggga gattccaagt 1020

gaagttgatc tgtatagcag tggtgagaac ggtgacgagt ctcctatagt taaggagcta 1080

cctaagatca gaacgcttcg tcgtgtggga gggagccctc gtgattcagg tgctgctaag 1140

agagttgcag aattacaagc caaggatcgt aacttcactt tttgccagct tctgaagttt 1200

gaagaatgaa tgatccgctt atcaatttga gtaaaatcca caactcttgt tgtggtt 1257

<210> 12

<211> 402

<212> PRT

<213> Arabidospis thaliana

<220>

<221> MISC_FEATURE

<223> seedy1 protein

<400> 12

Met Thr Ser Ile Glu Ala Thr Glu Thr Leu Asn Ala Pro Pro Lys Leu

1 5 10 15

Gln Ile Trp Asn Asn Ala Ala Phe Asp Asp Gly Asp Ser Gln Ile Thr
20 25 30

Ser Ala Ile Glu Ala Ser Ser Trp Ser His Leu Asn Glu Ser Phe Asp
35 40 45

Ser Asp Cys Ser Lys Glu Asn Gln Phe Pro Ile Ser Val Ser Ser Ser
50 55 60

Leu Gln Ser Ser Val Ser Ile Thr Glu Ala Pro Ser Ala Lys Ser Lys
65 70 75 80

Thr Val Lys Thr Lys Ser Ala Ala Asp Arg Ser Lys Lys Arg Asp Ile
85 90 95

Asp Ala Glu Ile Glu Glu Val Glu Lys Glu Ile Gly Arg Leu Ser Thr
100 105 110

Lys Leu Glu Ser Leu Arg Leu Glu Lys Ala Glu Gln Thr Ala Arg Ser
115 120 125

Ile Ala Ile Arg Gly Arg Ile Val Pro Ala Lys Phe Met Glu Ser Ser
130 135 140

Gln Lys Gln Val Lys Phe Asp Asp Ser Cys Phe Thr Gly Ser Lys Ser
145 150 155 160

Arg Ala Thr Arg Arg Gly Val Ser Leu Gly Pro Ala Glu Ile Phe Asn
165 170 175

Ser Ala Lys Lys Ser Glu Thr Val Thr Pro Leu Gln Ser Ala Gln Asn
180 185 190

Arg Arg Lys Ser Cys Phe Phe Lys Leu Pro Gly Ile Glu Glu Gly Gln
195 200 205

Val Thr Thr Arg Gly Lys Gly Arg Thr Ser Leu Ser Leu Ser Pro Arg
210 215 220

Ser Arg Lys Ala Lys Met Thr Ala Ala Gln Lys Gln Ala Ala Thr Thr
225 230 235 240

Val Gly Ser Lys Arg Ala Val Lys Lys Glu Glu Gly Val Leu Leu Thr
245 250 255

Ile Gln Pro Lys Arg Leu Phe Lys Glu Asp Glu Lys Asn Val Ser Leu
260 265 270

Arg Lys Pro Leu Lys Pro Gly Arg Val Val Ala Ser Arg Tyr Ser Gln
275 280 285

Met Gly Lys Thr Gln Thr Gly Glu Lys Asp Val Arg Lys Arg Ser Leu
290 295 300

Pro Glu Asp Glu Glu Lys Glu Asn His Lys Arg Ser Glu Lys Arg Arg
305 310 315 320

Ala Ser Asp Glu Ser Asn Lys Ser Glu Gly Arg Val Lys Lys Arg Trp
325 330 335

Glu Ile Pro Ser Glu Val Asp Leu Tyr Ser Ser Gly Glu Asn Gly Asp
340 345 350

Glu Ser Pro Ile Val Lys Glu Leu Pro Lys Ile Arg Thr Leu Arg Arg
355 360 365

Val Gly Gly Ser Pro Arg Asp Ser Gly Ala Ala Lys Arg Val Ala Glu
370 375 380

Leu Gln Ala Lys Asp Arg Asn Phe Thr Phe Cys Gln Leu Leu Lys Phe
385 390 395 400

Glu Glu

<210> 13

<211> 3074

<212> DNA

<213> Artificial sequence

<220>

<223> Sequence of the [PRO0090 - CDS0689 - terminator] expression casse
tte

<400> 13
cttctacatc ggcttaggtg tagcaacacg actttattat tattattatt attattatta 60
ttattttaca aaaatataaa atagatcagt ccctcaccac aagtagagca agttggtgag 120
ttattgtaaa gttctacaaa gctaatttaa aagttattgc attaacttat ttcatattac 180
aaacaagagt gtcaatggaa caatgaaaac catatgacat actataattt tgtttttatt 240
attgaaatta tataattcaa agagaataaa tccacatagc cgtaaagttc tacatgtggt 300
gcattaccaa aatatatata gcttacaaaa catgacaagc ttagtttgaa aaattgcaat 360
ccttatcaca ttgacacata aagtgagtga tgagtcataa tattattttc tttgctaccc 420
atcatgtata tatgatagcc acaaagttac tttgatgatg atatcaaaga acatttttag 480
gtgcacctaa cagaatatcc aaataatatg actcacttag atcataatag agcatcaagt 540
aaaactaaca ctctaaagca accgatggga aagcatctat aaatagacaa gcacaatgaa 600
aatcctcatc atccttcacc acaattcaaa tattatagtt gaagcatagt agtaatttaa 660
atcaactagg gatatcacaa gtttgtacaa aaaagcaggc tggtaccggt ccggaattcc 720
cgggatatcg tcgacccacg cgtccgctga cgcgtgggtt ccactacatc aagacatcta 780
ctacactcat ctttttgca cttattgggt gtaaattttt gaaacccagt tgagaaaaat 840
gagtgtgtta caatacccag aagggattga cccagcagat gttcagatat ggaacaatgc 900
agcatttgat aatggagatt ctgaagattt gtcttcgctg aaacgttctt ggtctcctct 960
gaaacccctt tcggttaggc catcagattc ctttgaatct gatttgtcaa gtaaggaaaa 1020
tcaaactcct ttatttgaga attcatctgt taatctctca tctccgttac ccataaagcc 1080
acttaaccct aatggggctc tggaaaattc aagactcaag ccgaacaagc ccaattccaa 1140
acagagtctt gatgagatgg cggctagaaa gagcggaaag ggaaatgatt tccgtgatga 1200
gaagaaaata gacgaggaaa ttgaagaaat tcagatggag attagtaggt tgagttcaag 1260
attagaggct ttgagaattg aaaaggctga gaaaactgtt gctaagactg ttgaaaagcg 1320
aggaagggtt gtggcagcaa agtttatgga gccaaaacaa agtgttatta agattgaaga 1380
gcgtatatca atgagtgcaa gaacaaaggt ggagcagaga aggggtctta gtttaggacc 1440
atctgagatt tttactggaa cgcggcggcg agggttgagt atggggccat cagatattct 1500
agcagggaca acaaaggcac ggcaattggg aaagcaagag atgattatta ctcctattca 1560
gccaatacaa aacaggcgaa agtcgtgttt ttggaagctt caagagattg aagaagaggg 1620
aaaaagttca agccttagtc ctaaatcaag aaaaactgct gcaagaacaa tggttacaac 1680
aaggcaggca gttactacaa ttgcatcaaa gaagaatttg aaaaaagatg atggactttt 1740

```
gagttcagtt cagccaaaga agttgtttaa agatctcgaa aagtctgctg ctgctaataa    1800
gaagccccag aggccgggga gggttgtggc tagtaggtat aatcagagta caattcagtc    1860
atcagtagtg agaaagaggt ctttacctga aaatgataag gatgagagta agagaaatga    1920
taagaaacgg tcgttatctg tagggaaaac gcgtgtgtct caaactgaga gcaagaattt    1980
gggtactgaa agtagggtga aaaagagatg ggaaattcct agtgagattg tagttcatgg    2040
aaacacagag agtgagaaat ctccactaag cattattgtg aagcctgatt tgcttccgcg    2100
aattaggatt gctcggtgtg tgaatgagac tcttagggat tctggacctg ctaaaagaat    2160
gatagagttg ataggcaaga aatcgttttt cagtagtgat gaagataagg agccacctgt    2220
ctgtcaagtt ttaagttttg cagaggaaga tgctgaagag gaataatgtg taataaaggg    2280
agctgctaac tcttttcatg ctctttcaat tttcaatcct gccttttaat ttttgttcat    2340
tcgtgccttt taattgaatg gggaagcatt cttttgcttc ctcaaactgg tattctagct    2400
tctgaattac attgtatggt acaatatgaa taaggttttg tcttccggca ggttgtccaa    2460
gttagttttt agcttaaaat agatgcggca gcggccgctc tagagtatcc ctcgaggggc    2520
ccaagcttac gcgtacccag ctttcttgta caaagtggtg atatcacaag cccgggcggt    2580
cttctaggga taacagggta attatatccc tctagatcac aagcccgggc ggtcttctac    2640
gatgattgag taataatgtg tcacgcatca ccatgggtgg cagtgtcagt gtgagcaatg    2700
acctgaatga acaattgaaa tgaaaagaaa aaaagtactc catctgttcc aaattaaaat    2760
tcattttaac cttttaatag gtttatacaa taattgatat atgttttctg tatatgtcta    2820
atttgttatc atccgggcgg tcttctaggg ataacagggt aattatatcc ctctagacaa    2880
cacacaacaa ataagagaaa aaacaaataa tattaatttg agaatgaaca aaaggaccat    2940
atcattcatt aactcttctc catccatttc catttcacag ttcgatagcg aaaaccgaat    3000
aaaaaacaca gtaaattaca agcacaacaa atggtacaag aaaaacagtt ttcccaatgc    3060
cataatactc gaac                                                      3074
```

<210> 14

<211> 668

<212> DNA

<213> Oryza sativa

<220>

<221> misc_feature

<223> prolamin RP6 promter sequence

<400> 14
ccttctacat cggcttaggt gtagcaacac gactttatta ttattattat tattattatt 60

attattttac aaaaatataa aatagatcag tccctcacca caagtagagc aagttggtga 120

gttattgtaa agttctacaa agctaattta aaagttattg cattaactta tttcatatta 180

caaacaagag tgtcaatgga acaatgaaaa ccatatgaca tactataatt ttgtttttat 240

tattgaaatt atataattca aagagaataa atccacatag ccgtaaagtt ctacatgtgg 300

tgcattacca aaatatatat agcttacaaa acatgacaag cttagtttga aaaattgcaa 360

tccttatcac attgacacat aaagtgagtg atgagtcata atattatttt tcttgctacc 420

catcatgtat atatgatagc cacaaagtta ctttgatgat gatatcaaag aacattttta 480

ggtgcaccta acagaatatc caaataatat gactcactta gatcataata gagcatcaag 540

taaaactaac actctaaagc aaccgatggg aaagcatcta taaatagaca agcacaatga 600

aaatcctcat catccttcac cacaattcaa atattatagt tgaagcatag tagtagaatc 660

caacaaca 668

<210> 15

<211> 7

<212> PRT

<213> Artificial sequence

<220>

<223> Motif 1 CORE SEQUENCE

<220>

<221> MISC_FEATURE

<222> (2)..(2)

<223> Xaa can be any amino acid

<220>

<221> MISC_FEATURE

<222> (5)..(6)

<223> Xaa can be any amino acid

<400> 15

Trp Xaa Asn Ala Xaa Xaa Asp
1   5

<210> 16

<211> 6

<212> PRT

<213> Artificial sequence

<220>

<223> Motif 2 CORE SEQUENCE

<220>

<221> MISC_FEATURE

<222> (4)..(5)

<223> Xaa can be any amino acid

<400> 16

Lys Glu Asn Xaa Xaa Pro
1   5

<210> 17

<211> 15

<212> PRT

<213> Artificial sequence

<220>

<223> Motif 3 (coiled coil) CORE SEQUENCE

<220>

<221> MISC_FEATURE

<222> (2)..(2)

<223> Xaa can be a stretch of 1 to 6 amino acids

<220>

<221> MISC_FEATURE

<222> (4)..(5)

<223> Xaa can be any amino acid

<220>

<221> MISC_FEATURE

<222> (8)..(10)

<223> Xaa can be any amino acid

<220>

<221> MISC_FEATURE

<222> (12)..(13)

<223> Xaa can be any amino acid

<400> 17

```
Glu Xaa Glu Xaa Xaa Arg Leu Xaa Xaa Xaa Leu Xaa Xaa Leu Arg
1               5                   10                  15
```

<210> 18

<211> 15

<212> PRT

<213> Artificial sequence

<220>

<223> Motif 4 CORE SEQUENCE

<220>

<221> MISC_FEATURE

<222> (3)..(3)

<223> Xaa can be any amino acid

<220>

<221> MISC_FEATURE

<222> (5)..(5)

<223> Xaa can be a stretch of 1 to 10 amino acids

<220>

<221> MISC_FEATURE

<222> (10)..(11)

<223> Xaa can be any amino acid

<220>

<221> MISC_FEATURE

<222> (14)..(14)

<223> Xaa can be a stretch of 1 to 6 amino acids

<400> 18

```
Leu Pro Xaa Ile Xaa Arg Asp Ser Gly Xaa Xaa Lys Arg Xaa Lys
1               5                   10                  15
```

CCR
PCT/EP2004/053030